Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **422,** 2854–2877 (2012) doi:10.1111/j.1365-2966.2012.20604.x

Constraining primordial non-Gaussianity with future galaxy surveys

Tommaso Giannantonio,[1,2]★ Cristiano Porciani,[2] Julien Carron,[3] Adam Amara[3] and Annalisa Pillepich[3,4]

[1]*Excellence Cluster Universe, Technical University Munich, Boltzmannstraße 2, D-85748 Garching bei München, Germany*
[2]*Argelander-Institut für Astronomie der Universität Bonn, Auf dem Hügel 71, D-53121 Bonn, Germany*
[3]*Institute for Astronomy, ETH Zurich, CH-8093 Zurich, Switzerland*
[4]*UCO/Lick Observatory, Department of Astronomy and Astrophysics, University of California Santa Cruz, Santa Cruz, CA 95064, USA*

Accepted 2012 January 19. Received 2012 January 18; in original form 2011 September 5

ABSTRACT

We study the constraining power on primordial non-Gaussianity of future surveys of the large-scale structure of the Universe for both near-term surveys (such as the Dark Energy Survey – DES) as well as longer term projects such as *Euclid* and *WFIRST*. Specifically we perform a Fisher matrix analysis forecast for such surveys, using DES-like and *Euclid*-like configurations as examples, and take account of any expected photometric and spectroscopic data. We focus on two-point statistics and consider three observables: the 3D galaxy power spectrum in redshift space, the angular galaxy power spectrum and the projected weak-lensing shear power spectrum. We study the effects of adding a few extra parameters to the basic Λ cold dark matter (ΛCDM) set. We include the two standard parameters to model the current value for the dark-energy equation of state and its time derivative, w_0, w_a, and we account for the possibility of primordial non-Gaussianity of the local, equilateral and orthogonal types, of parameter $f_{\rm NL}$ and, optionally, of spectral index $n_{f_{\rm NL}}$. We present forecasted constraints on these parameters using the different observational probes. We show that accounting for models that include primordial non-Gaussianity does not degrade the constraint on the standard ΛCDM set nor on the dark-energy equation of state. By combining the weak-lensing data and the information on projected galaxy clustering, consistently including all two-point functions and their covariance, we find forecasted marginalized errors $\sigma(f_{\rm NL}) \sim 3$, $\sigma(n_{f_{\rm NL}}) \sim 0.12$ from a *Euclid*-like survey for the local shape of primordial non-Gaussianity, while the orthogonal and equilateral constraints are weakened for the galaxy clustering case, due to the weaker scale dependence of the bias. In the lensing case, the constraints remain instead similar in all configurations.

Key words: cosmological parameters – early Universe – large-scale structure of the Universe.

1 INTRODUCTION

The accurate measurement of the cosmic microwave background (CMB) anisotropies by the *Wilkinson Microwave Anisotropy Probe* (*WMAP*) satellite (Larson et al. 2011), together with the observation of Type Ia supernovae (SNe) and the deep and wide mapping of the large-scale structure (LSS) of the Universe, as done by the Sloan Digital Sky Survey (SDSS) survey (Aihara et al. 2011), has led to the standard model of cosmology, whose parameters are now measured with unprecedented accuracy. However, the simplest model that fits all observations, Λ cold dark matter (ΛCDM), has some shortcomings and poorly understood phases, in particular the late-time acceleration (or dark energy) (Frieman, Turner & Huterer 2008) and the early-time inflationary phase (Linde 2008). For these reasons, it was agreed that new generations of surveys should be built to improve upon the current data. The group known as Dark Energy Task Force (DETF; Albrecht et al. 2006) has then labelled current finished and ongoing surveys as stages I and II, while future experiments belong to stage III or IV depending on their time-scale and forecasted power, such as the Dark Energy Survey (DES; Abbott et al. 2005) and the *Euclid* satellite (Laureijs 2009; Laureijs et al. 2011), respectively.

These surveys will also shed new light on other basic questions, such as the origin of the LSS and the physics of the early universe. It is currently assumed that the observed inhomogeneities in the matter density were seeded by quantum fluctuations at primordial times. These fluctuations were then stretched by inflation (or some

★E-mail: tommaso.giannantonio@universe-cluster.de

alternative theory) on superhorizon scales and acted as seeds first for the observed CMB anisotropies, and then for the matter density fluctuations. The simplest possible model for the primordial fluctuations consists of a Gaussian random field with a nearly scale-invariant power spectrum (Bardeen et al. 1986). However, many models exist for the inflationary phase (Lyth & Liddle 2009), many of which would alter the statistics of the fluctuations, in particular introducing a non-Gaussian component (Chen 2010). It is therefore important to put observational constraints on the presence (and, once detected, on the amount) of primordial non-Gaussianity (PNG), in order to distinguish between models of the early universe. This is usually done in terms of the bispectrum amplitude $f_{\rm NL}$ for a set of template shapes for the bispectrum of the perturbations (Liguori et al. 2010, see Appendix A for a more precise definition).

The most well-established method to constrain PNG is to measure the three-point statistics of the CMB temperature anisotropies (Komatsu 2010). Current limits from *WMAP*7 for the local, equilateral and orthogonal types of PNG are $-10 < f_{\rm NL}^{\rm loc} < 74$, $-214 < f_{\rm NL}^{\rm equ} < 266$ and $-410 < f_{\rm NL}^{\rm ort} < 6$ at 95 per cent CL, respectively (Komatsu et al. 2011). This will be greatly improved by the *Planck* mission, for which the expected uncertainty is of the order of $\sigma(f_{\rm NL}^{\rm loc}) \simeq 5$ (Komatsu & Spergel 2001), limited by cosmic variance. A complementary way is to constrain PNG from LSS studies. In this case, one can exploit multiple effects. First, the abundance of the most massive structures existing at any given cosmic time strongly depends on the level of PNG (Lucchin & Matarrese 1988; Pillepich, Porciani & Hahn 2010) although the effect is degenerate with other cosmological parameters (Pillepich, Porciani & Reiprich 2011). The recent detection of a few high-redshift clusters gives some evidence in favour of non-Gaussian models (Hoyle, Jimenez & Verde 2011), even though no consensus has been reached (Mortonson, Hu & Huterer 2011; Hoyle et al. 2012) also due to the low-number statistics. The abundance of weak-lensing peaks has also been proposed as a method to measure PNG (Marian et al. 2011). Moreover, in full analogy with the CMB case, any three-point statistic of the galaxy distribution is affected by PNG (Sefusatti & Komatsu 2007; Jeong & Komatsu 2009). However, the situation is complicated by the overlapping effects of primordial and late-time non-Gaussianity, which is driven by the non-linear growth of structure and galaxy biasing (Komatsu et al. 2009). These effects can be disentangled only by a thorough understanding of the non-linear regime. Similar considerations apply to modifications of the matter power spectrum on small scales induced by PNG (Taruya, Koyama & Matsubara 2008; Pillepich, Porciani & Hahn 2010). This is why the recent discovery that PNG also generates a strong scale dependence of galaxy biasing on very large scales (Dalal et al. 2008) has attracted particular interest in the literature. It has been shown that measurements of PNG based on this single feature are currently competitive with CMB studies (Slosar et al. 2008; Xia et al. 2011) and will remain so in the future (Afshordi & Tolley 2008; Carbone, Mena & Verde 2010).

In this paper we want to investigate further these techniques, applying the Fisher matrix formalism to determine to what accuracy we expect future surveys will constrain PNG using two-point statistics only. We will compare the constraints from the large- and small-scale effects of the LSS with those expected by the CMB *Planck* mission, extending the results by Joachimi & Bridle (2010), Fedeli & Moscardini (2010), Carbone et al. (2010), Wang et al. (2010) and Namikawa, Okamura & Taruya (2011). In particular, our study differs from the previous ones as the forecasts on PNG by Fedeli & Moscardini (2010) only considered weak-lensing data and did not include variations on any other cosmological parameter but $f_{\rm NL}$; the results by Carbone et al. (2010) only considered the 3D galaxy power spectrum and only the local type of PNG, while Joachimi & Bridle (2010) and Wang et al. (2010) did not consider PNG.

The plan of the paper is as follows: we will review the effects of PNG on the LSS in Section 2, while in Section 3, we will describe how we calculate the non-linear power spectra (for galaxy clustering and weak-lensing studies) in terms of the cosmological parameters. We will then describe the future surveys considered in Section 4, present our basic forecasts in Section 5 and expand on them in Section 6. We shall finally conclude in Section 7.

2 THE EFFECTS OF PRIMORDIAL NON-GAUSSIANITY

2.1 Definitions

The simplest single-field, slow-roll model for inflation gives rise to a nearly Gaussian distribution of the curvature perturbations ζ or the Bardeen potential Φ at primordial times corresponding to a redshift z_*. This changes however in most generalizations: many models, and especially multi-field inflation, produce non-Gaussianities (see e.g. the recent review by Byrnes & Choi 2010). There are different possibilities for a departure from a purely Gaussian distribution. The most general expression of a deviation from Gaussianity at a quadratic level can be expressed by a non-local relationship between the primordial Bardeen's potential Φ and a Gaussian auxiliary potential φ. In real space (Schmidt & Kamionkowski 2010)

$$\Phi(\boldsymbol{x}, z_*) = \varphi(\boldsymbol{x}, z_*) + (f_{\rm NL} * W * \varphi * \varphi)(\boldsymbol{x}, z_*), \tag{1}$$

where the asterisk denotes convolution and $W(\boldsymbol{y}, \boldsymbol{z})$ is a kernel whose form describes the type of non-Gaussianity considered. This is often called 'the CMB definition', as opposed to 'the LSS definition', as it is written at early times z_*. Here the $f_{\rm NL}$ function quantifies the amount of PNG at first order, and it represents the first relevant deviation to measure, i.e. the skewness of the perturbations' probability distribution at a given length scale. Subsequent contributions at higher order, e.g., the kurtosis contribution $g_{\rm NL}$, are also expected in many theories, but will not be considered in the following. In the simplest models, $f_{\rm NL}$ does not depend on scale, and thus it simply assumes a constant value in equation (1). In most of the following we will assume this simplification, but we will extend the analysis to the scale-dependent case in Section 6.5.

We can define the power spectrum and bispectrum of the potential in the usual way as

$$\begin{aligned} \langle \tilde{\Phi}(k)\, \tilde{\Phi}(k') \rangle &= (2\pi)^3\, \delta_D(k + k')\, P_\Phi(k) \\ \langle \tilde{\Phi}(k)\, \tilde{\Phi}(k')\, \tilde{\Phi}(k'') \rangle &= (2\pi)^3\, \delta_D(k + k' + k'')\, B_\Phi(k, k', k''), \end{aligned} \tag{2}$$

where the tilde denotes Fourier transformation. Neglecting subdominant corrections from the four-point correlator (trispectrum), we can also assume $P_\Phi(k) \simeq P_\varphi(k)$ at leading order in $f_{\rm NL}$. Then it can be seen that applying the definition of equation (1), the kernel W defines the relationship between the power spectrum and bispectrum as

$$B_\Phi(k, k', k'') = 2 f_{\rm NL} \left[\tilde{W}(k, k')\, P_\Phi(k)\, P_\Phi(k') + 2\,\text{perms.} \right], \tag{3}$$

assuming constant $f_{\rm NL}$. In the simplest case of PNG of the local type, $W = 1$ and the bispectrum peaks for squeezed triangles ($k \ll k' \sim k''$); besides this case, we will consider in the following also

the equilateral configuration, for which the bispectrum is maximum for $k \sim k' \sim k''$, and the orthogonal type, which was constructed to be orthogonal to the previous two types. These configurations or a mixture thereof can be produced under different inflationary scenarios; see e.g. Babich, Creminelli & Zaldarriaga (2004) and Senatore, Smith & Zaldarriaga (2010) for more details. The bispectra for the three cases are given in Appendix A.

The analysis of CMB bispectra from the *Planck* satellite is currently considered to be the most promising tool for distinguishing between these scenarios (Fergusson & Shellard 2009). In this paper we will consider the three configurations separately, although an analysis of a general linear combination of the three modes is in principle possible.

The primordial Bardeen potential (equal to the gravitational potential with the opposite sign for subhorizon modes) is then related by the Poisson equation to the total matter density perturbations δ. At a linear level

$$\tilde{\delta}(\boldsymbol{k}, z) = \alpha(k, z)\, \tilde{\Phi}(\boldsymbol{k}, z_*) \simeq \alpha(k, z)\, \tilde{\varphi}(\boldsymbol{k}, z_*)\,, \tag{4}$$

where

$$\alpha(k, z) = \frac{2\, c^2\, k^2\, T(k)\, D(z)}{3\Omega_{\rm m} H_0^2}\, \frac{g(0)}{g(z_*)}\,, \tag{5}$$

and we have introduced the linear growth function $D(z)$ (normalized so that $D(0) = 1$), the transfer function $T(k)$ and the potential growth function $g(z) \propto (1 + z)\, D(z)$. We can therefore write the tree-level matter power spectrum,

$$P_0(k, z) = \alpha^2(k, z) P_\Phi(k, z_*) \simeq \alpha^2(k, z) P_\varphi(k, z_*). \tag{6}$$

2.2 LSS and primordial non-Gaussianity

The LSS of the Universe is commonly described in terms of different tracers. In high-density regions of the underlying dark-matter density contrast δ, we observe the formation of dark-matter haloes, whose distribution is described by the field $\delta_{\rm h}$. We expect that galaxy formation occurs within these haloes, thus producing a galaxy density field $\delta_{\rm g}$. On very large scales and to first approximation, the linear bias coefficients $b_{\rm h}$ and $b_{\rm g}$ relate the different density fields as $\delta_{\rm h} \simeq b_{\rm h}\, \delta$ and $\delta_{\rm g} \simeq b_{\rm g}\, \delta$. From the observational side, the dark-matter distribution δ can be directly observed with weak-lensing studies, while $\delta_{\rm g}$ can be measured by mapping the galaxy distribution. The most natural two-point statistic that can be observed is then the power spectrum of these density fields which can be defined in analogy with equation (2). Therefore, the galaxy and matter power spectra approximately satisfy $P_{\rm g}(k) \simeq b_{\rm g}^2\, P_{\rm m}(k)$ and a similar relation can be written in terms of $b_{\rm h}$ (which depends on halo mass and time) for the dark-matter haloes.

The introduction of PNG has multiple observable consequences for the two-point statistics. First, a small modification of the matter power spectrum appears on small scales (Taruya et al. 2008; Pillepich et al. 2010), due to corrections coming from the linear matter bispectrum, which is non-vanishing in this case. Secondly, the biasing law between dark-matter haloes and the underlying mass-density field is altered, becoming strongly scale dependent in the local and orthogonal cases (Afshordi & Tolley 2008; Dalal et al. 2008; Matarrese & Verde 2008; Slosar et al. 2008; Giannantonio & Porciani 2010; Schmidt & Kamionkowski 2010; Valageas 2010; Desjacques, Jeong & Schmidt 2011). This is due to the coupling between long- and short-wavelength modes of the perturbations, affecting the halo power spectrum on large scales. In addition to this effect, there is also a smaller scale-independent modification to the bias coming from the non-Gaussian form of the mass functions (Slosar et al. 2008; Giannantonio & Porciani 2010), affecting the power spectrum on all scales. The above-mentioned numerous independent calculations and comparisons with N-body simulations have shown that the bias at fixed halo mass can be written as

$$b_{\rm eff}(k, f_{\rm NL}) = b(f_{\rm NL} = 0) + \delta b(f_{\rm NL}) + \Delta b(k, f_{\rm NL})\,, \tag{7}$$

where $\delta b(f_{\rm NL})$ and $\Delta b(k, f_{\rm NL})$ denote the scale-independent and the scale-dependent corrections, respectively. If we introduce a weighted variance smoothed by a top-hat filter $F_R(k)$,

$$\sigma_{R,n}^2 \equiv \int \frac{{\rm d}^3 k}{(2\pi)^3}\, k^n\, P_0(k)\, F_R^2(k)\,, \tag{8}$$

then the deviation at any given redshift for each configuration is approximately given by (Schmidt & Kamionkowski 2010; Desjacques et al. 2011)

$$\begin{aligned}
\Delta b^{\rm loc}(k, f_{\rm NL}) &\simeq \frac{2 f_{\rm NL}\, \delta_{\rm c}\, b_{\rm L}}{\alpha(k)} \\
\Delta b^{\rm ort}(k, f_{\rm NL}) &\simeq -\frac{6 f_{\rm NL}\, k}{\alpha(k)}\, \frac{\sigma_{R,-1}^2}{\sigma_{R,0}^2}\, \left[\delta_{\rm c}\, b_{\rm L} + 2\left(\frac{\sigma'_{R,-1}\, \sigma_{R,0}}{\sigma_{R,-1}\, \sigma'_{R,0}} - 1\right)\right] \\
\Delta b^{\rm equ}(k, f_{\rm NL}) &\simeq \frac{6 f_{\rm NL}\, k^2}{\alpha(k)}\, \frac{\sigma_{R,-2}^2}{\sigma_{R,0}^2}\, \left[\delta_{\rm c}\, b_{\rm L} + 2\left(\frac{\sigma'_{R,-2}\, \sigma_{R,0}}{\sigma_{R,-2}\, \sigma'_{R,0}} - 1\right)\right],
\end{aligned} \tag{9}$$

where the prime denotes derivatives w.r.t. R. Here the (linear) Lagrangian halo bias $b_{\rm L} = b(f_{\rm NL} = 0) + \delta b(f_{\rm NL}) - 1$ includes the (generally small) scale-independent PNG correction. In the spirit of the halo model, we will assume that the galaxy linear bias coefficient $b_{\rm g}$ is given, on large scales, by a weighted average of the halo bias at different masses. Note that, on scales larger than $\sim$1 Mpc, which we consider only in our analysis, the scale dependence is translated unaffected from the halo to the galaxy bias. In detail, for any choice of galaxy number density $n(M)$ and halo occupation distribution (HoD) of mean number $N_{\rm g}(M)$, we can write

$$b_{\rm g}(k) \propto \int b_{\rm h}(k, M)\, n(M)\, N_{\rm g}(M)\, {\rm d}M\,, \tag{10}$$

which means that, independently from the HoD model, the behaviour of the galaxy bias is the same as for the halo bias:

$$b_{\rm g}(k, f_{\rm NL}) = b_{\rm g}(f_{\rm NL} = 0) + \delta b_{\rm g}(f_{\rm NL}) + \Delta b_{\rm g}(k, f_{\rm NL}). \tag{11}$$

Therefore, when dealing with galaxy clustering, we will rewrite the Lagrangian bias as $b_{\rm L} = b - 1$ with b a nuisance parameter to be marginalized over (the index g is understood). From the expressions above, we can see that the asymptotical scale dependence is $\propto k^{-2}$, k^{-1}, k^0 for the three configurations, respectively.

2.2.1 *Fudge factors*

Finally, a note of caution. Equations (9) have been derived using the peak-background split technique to the halo mass function obtained through a Press–Schechter ansatz, i.e. assuming that linear density perturbations on a given mass scale collapse into dark-matter haloes if $\delta > \delta_{\rm c}$. To first approximation, it is often assumed that $\delta_{\rm c} \simeq 1.686$ (independent of halo mass) as expected from the spherical collapse model in an Einstein–de Sitter universe. However, more rigorous applications of the excursion-set approach show that the correction to the halo mass function for PNG is more complicated than this. Using models of triaxial collapse, Lam & Sheth (2009) found that the critical threshold is mass dependent and tends to $\sqrt{0.7}\delta_{\rm c}$ for large halo masses (where the factor 0.7 is determined by fitting numerical

simulations). Alternatively, accounting for the fluctuations of the collapse threshold measured in N-body simulations by Robertson et al. (2009) gives an effective threshold $0.89\delta_c$ (Maggiore & Riotto 2010a,b).

Inspired by this theoretical work, in numerical applications, the standard collapse threshold δ_c is sometimes rescaled by a fudge factor q; this is a coefficient of order unity which is introduced heuristically to improve the agreement of the analytical models for the halo mass function and bias with N-body simulations (Carbone, Verde & Matarrese 2008; Grossi et al. 2009; Desjacques & Seljak 2010b; Pillepich et al. 2010; Wagner & Verde 2011). However, there are several ambiguities concerning this factor. First, there is no consensus on the actual role of q. Some authors use it in the ratio (equation A10 in Appendix A) between the halo mass functions derived from Gaussian and non-Gaussian initial conditions (Giannantonio & Porciani 2010; Pillepich et al. 2010). Using the peak-background split to derive the halo bias, this gives equation (9) with the replacement $\delta_c \to q\delta_c$ (Giannantonio & Porciani 2010). On the other hand, other authors use the correction $\sqrt{q}\delta_c$ in the halo mass function and $q\delta_c$ in the halo bias (Grossi et al. 2009). Furthermore, it is sometimes assumed that there are two independent fudge factors for the mass functions and the bias (Wagner & Verde 2011). Secondly, even assuming that q is not mass and redshift dependent, significantly different values of q (with nearly 30 per cent deviations) are required to reproduce the N-body results when dark-matter haloes in the simulations are identified using either the friends-of-friends (FoF) or the spherical overdensity (SO) method: this makes unclear which value should actually be used for galaxies (Desjacques & Seljak 2010a). Finally, the improved model by Desjacques et al. (2011) that we adopt shows good agreement with N-body simulations without the need for any fudge factor, at least for a specific choice of the halo finder. Because of these reasons, and because the theoretical motivation favouring a specific value of q is still lacking, in this analysis we will set $q = 1$ throughout. This choice will affect our models for the non-linear mass power spectrum (through the halo mass function in the halo model) and for the galaxy power spectrum (through both the halo mass function and the scale dependence of the galaxy bias coefficient). The impact of q on the weak-lensing observables is subtle and somewhat degenerate with other cosmological parameters. In fact, in equation (A10) $f_{\rm NL}$ is multiplied by a polynomial in q whose coefficients depend on the mass variance. On the other hand, galaxy clustering studies will constrain the amplitude of PNG almost entirely from the scale-dependent bias on the largest scales. This implies that observations will basically constrain the product $qf_{\rm NL}$ and, at least in this case, it will be trivial to correct our results whenever future progress on the fudge factor q will be made. Another possibility would be to use q as a nuisance parameter to be marginalized over after assuming a theoretically based prior. Even though we will not consider this option in this paper, it is clear that this marginalization would degrade the constraints on $f_{\rm NL}$ by a similar amount to the uncertainty in q.

3 THE NON-LINEAR REGIME

In order to take advantage of the extended range of measurements which will be available from future surveys, it is desirable to extend the analyses as far as possible into the non-linear regime. For the galaxy clustering, the presence of non-linear biasing complicates the issue, so to be conservative, it is reasonable to expect an accurate theoretical modelling only up to scales $k_{\rm max} \simeq 0.15\,h\,{\rm Mpc}^{-1}$ at $z = 0$, even though new theoretical methods are being proposed which may extend the accuracy of the predictions to smaller scales, such as e.g. by Simpson et al. (2011). However, in the case of weak lensing, the calculations are made simpler by the absence of the bias, and a reasonably accurate modelling is possible up to the smallest scales, if the effect of baryons can be neglected or accurately modelled [see e.g. van Daalen et al. (2011) for a quantitative analysis of this issue]. As in most forecast papers appeared so far we will not consider this effect in detail here.

3.1 Halo model

We compute the non-linear mass power spectrum using the halo model developed by Ma & Fry (2000) and Seljak (2000) based on an original idea by McClelland & Silk (1977). See Cooray & Sheth (2002) for a review. In this approach all the matter in the Universe is assumed to be concentrated in a set of discrete and clustered dark-matter haloes. The non-linear power spectrum can be written as a sum of two terms:

$$P_{\rm m}(k, z, f_{\rm NL}) = P_1(k, z, f_{\rm NL}) + P_2(k, z, f_{\rm NL}). \tag{12}$$

The two-halo term P_2 dominates on large scales and represents the correlations between mass concentrations lying in different haloes, while the one-halo term P_1 describes correlations of particles which belong to the same halo. These two terms can be calculated as

$$\begin{aligned} P_1(k, z, f_{\rm NL}) &= \int_0^\infty n(M, z, f_{\rm NL}) \left[\frac{\tilde{\rho}(M, z, k)}{\rho_{\rm m}}\right]^2 {\rm d}M \\ P_2(k, z, f_{\rm NL}) &\\ &= \left[\int_0^\infty n(M, z, f_{\rm NL})\, b(M, z, k, f_{\rm NL})\, \frac{\tilde{\rho}(M, z, k)}{\rho_{\rm m}} {\rm d}M\right]^2 P_0(k, z), \end{aligned} \tag{13}$$

and there are three ingredients which are needed to implement this model: the halo mass function n, the linear halo bias b and the halo density profile ρ of the mean value $\rho_{\rm m}$. The tree-level matter power spectrum P_0 is defined in equation (6); it depends on cosmology through its slope, the growth function and the transfer function, and it does not include any $f_{\rm NL}$ correction. There are many possible ways to parametrize these ingredients, and we have explored different alternatives trying to maximize the agreement with the power spectra measured from the N-body simulations presented in Pillepich et al. (2010) (hereafter PPH08). Concerning the Gaussian part of the halo mass function, we have tried the fitting formula by Tinker et al. (2008), which is based on dark-matter haloes identified with an SO halo finder in N-body simulations, and the similar function by PPH08, which makes use of an FoF halo finder. For the value of the mean overdensity $\Delta_{\rm V}$ which is enclosed by the virial radius of haloes, we tried the redshift scaling relation by Bryan & Norman (1998) and a constant value of $\Delta_{\rm V} = 200$. We also explored the halo exclusion prescriptions (Zheng et al. 2002; Magliocchetti & Porciani 2003; Tinker et al. 2005). We found that the best agreement with the simulations was recovered using the mass function by PPH08 and $\Delta_{\rm V} = 200$, which we will use in the analysis. The agreement is at the $\sim$10 per cent level as can be seen in Fig. 1. For non-Gaussian models, we will then apply the appropriate corrections to the PPH08 halo mass function using the method developed by LoVerde et al. (2008) and summarized in equation (A10). We will consistently derive the non-Gaussian halo bias using the peak-background split formalism, as reviewed in Appendix A.

We have assumed the analytical description by Navarro, Frenk & White (1996, 1997) for the halo density profile of dark-matter haloes (NFW). In this approach, the radial density ρ is written in real space as

$$\rho(r) = \frac{\rho_{\rm s}}{(r/r_{\rm s})\,(1 + r/r_{\rm s})^2}, \tag{14}$$

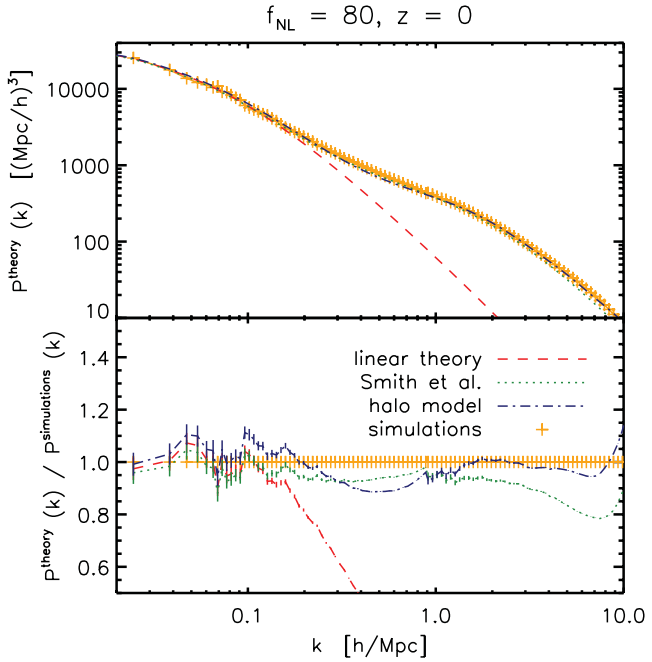

Figure 1. Comparison of the theoretical matter power spectra with the measurements from N-body simulations by PPH08, with local $f_{\rm NL} = 80$. The top panel shows the total $P_{\rm m}(k)$ for the linear (red) and two non-linear models: the halo model (blue) and the model by Smith, Scoccimarro & Sheth (2007) (green). Below we plot the ratio $P_{\rm m}(k)/P_{\rm simulations}(k)$. All curves are at $z = 0$: a similar (∼10 per cent) level of accuracy is obtained at other redshifts.

and it is fully specified by the parameters $r_{\rm s}$,$\rho_{\rm s}$, describing a scaling radius where the density profile changes its slope and its associated density. Alternatively, we can use the concentration $c \equiv R_{\rm V}/r_{\rm s}$, where the virial radius $R_{\rm V}$ is defined as the radius of the sphere whose mean density is $\Delta_{\rm V} = 200$ times the average density of the Universe. We can finally write the density profile in Fourier space as (Scoccimarro et al. 2001; Rudd, Zentner & Kravtsov 2008)

$$\tilde{\rho}(M,z,k) = 4\pi\rho_{\rm s}r_{\rm s}^3 \left\{ \sin(kr_{\rm s})\,[{\rm Si}(kr_{\rm s}(1+c)) - {\rm Si}(kr_{\rm s})] - \frac{\sin(ckr_{\rm s})}{kr_{\rm s}(1+c)} + \cos(kr_{\rm s})\,[{\rm Ci}(kr_{\rm s}(1+c)) - {\rm Ci}(kr_{\rm s})] \right\}. \quad (15)$$

Here Si(x), Ci(x) are the sine and cosine integrals, respectively. The concentration is known to depend on halo mass and redshift. We use the formula by Klypin, Trujillo-Gomez & Primack (2011), obtained from the Bolshoi simulation:

$$c(M,z) = 9.2\,\kappa(z)\,D^{1.3}(z) \left(\frac{M}{10^{12}\,h^{-1}\,{\rm M}_\odot}\right)^{-0.09} \times \left[1 + 0.013 \left(\frac{M}{10^{12}\,h^{-1}\,{\rm M}_\odot} D(z)^{-\frac{1.3}{0.09}}\right)^{0.25}\right], \quad (16)$$

where the function $\kappa(z)$ is introduced to correct for the different definition of the virial radius, and it has values $\kappa(z = 0) = 1.26$ and $\kappa(z \geq 1) \simeq 0.96$. We have also checked that using the result by Bullock et al. (2001) for the concentration does not change significantly the results. It is worth noting that the concentration is in principle dependent on the PNG parameter $f_{\rm NL}$ but, as was shown e.g. by Smith, Desjacques & Marian (2011) and D'Aloisio & Natarajan (2011), the effect is very small (at least for local PNG) and we will ignore it in the rest of this analysis.

3.2 Power spectrum and comparison with simulations

With the ingredients described above, we can now calculate the full non-linear matter power spectrum $P_{\rm m}(k, z)$; to recover the linear regime in the large-scale limit, it is fundamental that

$$\int_0^\infty n(M,z,f_{\rm NL})\,b(M,z,k,f_{\rm NL})\,\frac{M}{\rho_{\rm m}}{\rm d}M = 1 \qquad (k \to 0). \quad (17)$$

We will enforce this constraint in our calculations by adding a constant to the bias in the smallest mass bin, as also described by Fedeli & Moscardini (2010). This condition should be automatically satisfied, but the need to enforce it explicitly arises due to the binning, the finite integration limits which exist in practice and other numerical issues.

We then compare the results of the non-linear power spectrum obtained with these recipes with the N-body simulations by PPH08. Consistently with PPH08, for our calculations we assume the *WMAP*5 flat ΛCDM model, with parameters $h = 0.701$, $\sigma_8 = 0.817$, $n_{\rm s} = 0.96$, $\Omega_{\rm m} = 0.279$, $\Omega_{\rm b} = 0.0462$, $\Omega_\Lambda = 0.721$. We show the result of the comparisons in Fig. 1 at fixed redshift $z = 0$, as the comparison results are very similar at different redshifts. In the top panel we can see the full matter power spectra for the choice of local $f_{\rm NL} = 80$, compared with the simulations. Below we show the residuals $P_{\rm m}(k)/P_{\rm simulations}(k)$. We can see that the agreement is at the 10 per cent level up to scales of 10 h Mpc^{-1}. This is the level of accuracy to be expected from the halo model (Cooray & Sheth 2002). We can also see a comparison with the method to calculate the non-linear power spectrum using a fit from simulations by Smith et al. (2007), whose accuracy is of similar order for $k < 2\,h$ Mpc^{-1}, but then degrades further to the ∼20 per cent level.

In Section 6.8 we will compare the results of the Fisher matrix analysis obtained using different models for the non-linear evolution of the matter power spectrum, finding that the level of uncertainty still present in the halo model does not significantly impact the results of our forecasts in the galaxy clustering case.

4 FUTURE GALAXY SURVEYS

Wide galaxy surveys stand amongst the most powerful ways of probing cosmology; they can be used to measure several observables, such as tracers of the LSS of the Universe and indicators of the expansion rate. For example the SDSS (Aihara et al. 2011) and its follow-ups have greatly improved our understanding of the standard model, especially thanks to the measurement of galaxy clustering and baryon acoustic oscillations (see e.g. Percival et al. 2007), but also using galaxy clusters (Koester et al. 2007) and Type Ia SNe (Kessler et al. 2009; Lampeitl et al. 2009). Weak-lensing surveys (see e.g. Fu et al. 2008) are becoming increasingly powerful in constraining cosmology due to the improved control of systematics. The same surveys also provide additional cosmological probes, e.g., their external correlations with the CMB anisotropies (Giannantonio et al. 2008).

All of these observables share part of the information and will therefore not be fully independent. A considerable challenge for future surveys will be to properly account for their covariance, as was discussed e.g. for lensing and clustering by Hu & Jain (2004) and for lensing and galaxy clusters by Takada & Bridle (2007). In this paper we study how the constraining power on cosmology will improve in the near future using two examples of upcoming surveys, focusing on the DES (Abbott et al. 2005) for the DETF stage III and a *Euclid*-like survey (Laureijs 2009) for stage IV. It is reasonable to expect that other surveys in the same class will

Table 1. Technical specifications of the surveys used for forecasts.

Parameter	Description	*Euclid* photometric	*Euclid* spectroscopic	DES
$\sigma_z(z)/(1+z)$	Redshift uncertainty	0.05	0.001	0.1
$\bar{z}$	Median redshift	1.0	1.1	0.8
n	Galaxy density	40 arcmin^{-2}	0.7775 arcmin^{-2}	12 arcmin^{-2}
A	Surveyed area	20 000 square degrees	20 000 square degrees	5000 square degrees
γ	Intrinsic ellipticity noise	0.247	–	0.16
$\mathrm{d}N/\mathrm{d}z(z)$	Galaxy distribution	Smail et al.	Geach et al.	Smail et al.
	Parameters of galaxy distribution	$\alpha = 2$	Flux cut $= 4 \times 10^{-16}$ erg s^{-1} cm^{-2}	$\alpha = 2$
–	–	$\beta = 1.5$	Efficiency = 35 per cent	$\beta = 1.5$
M	Number of redshift bins	12	12	8

have comparable performances, such as in particular the American proposed *WFIRST* class IV mission.

4.1 Stage III: the Dark Energy Survey

The DES[1] (Abbott et al. 2005) is an optical and near-infrared (IR) survey which was deployed at the Cerro Tololo observatory in the Chilean Andes, and has recently come to light in late 2011. The wide-field survey will cover 5 000 square degrees in the Southern sky, reaching ∼24 mag in the SDSS bands g, r, i, z and Y. Additionally, the J, H, K bands for the same fields are expected to be added from the European Southern Observatory (ESO) Vista survey. About 300 million galaxies are expected to be observed with shapes, photo-z's and positions, as well as 100 000 galaxy clusters and 1000 Type Ia SNe using a smaller repeated imaging survey.

In this paper we will study the Fisher matrix forecasted errors on dark energy and PNG using the weak lensing, galaxy clustering and the combined data from this survey. For this purpose, we will assume the specifications summarized in Table 1 and that the redshift distribution of the sources can be well approximated by the law (Smail, Ellis & Fitchett 1994)

$$\frac{\mathrm{d}N}{\mathrm{d}z}(z) = \frac{1}{\Gamma\left(\frac{\alpha+1}{\beta}\right)}\,\beta\,\frac{z^{\alpha}}{z_0^{\alpha+1}}\,\exp\left[-\left(\frac{z}{z_0}\right)^{\beta}\right], \tag{18}$$

where the parameters are set to $\alpha = 2$, $\beta = 1.5$, in which case z_0 is related to the median redshift by $z_0 \simeq \bar{z}/1.41$. We will finally split this distribution into eight equally populated redshift bins. Due to the uncertainty in the photo-z determination, we will write the theoretical redshift distribution of the sources as the convolution of these redshift bins with a Gaussian characterized by a dispersion matching the photometric redshift uncertainty $\sigma_z(z)$. This approach assumes that the photometric redshift estimation is perfectly calibrated to a spectroscopic sample; the distribution of the redshift errors is Gaussian and the rms error $\sigma_z(z)$ is known; catastrophic errors can be identified and down-weighted in the analysis. See Kitching, Taylor & Heavens (2008) and Bernstein & Huterer (2010) for more detailed approaches which take into account more of the possible systematics arising in this case.

Finally, in the absence of a detailed model we will assume the fiducial value of the galaxy bias to follow the law $b(z) = \sqrt{1+z}$ following Rassat et al. (2008), but we will study the effect of modifying the fiducial bias in Section 6.4.

4.2 Stage IV: *Euclid*

Euclid[2] is a proposed space mission of the European Space Agency (ESA), which is expected to survey the whole extragalactic sky (up to 20 000 square degrees) from the L2 point in space (Laureijs 2009). Launch is currently scheduled in 2018. In our analysis we approximate this planned survey with the settings and specifications described in the *Euclid* Assessment Study Report (Laureijs 2009). Although the specifications of the survey have since evolved (Laureijs et al. 2011), this does not affect the current study because our intention is to draw broad conclusions that will be relevant to the whole class of stage IV surveys including *WFIRST*, *LSST* and others. See in any case Section 6.10 for a discussion of how the results change when using the latest Red Book specifications. The *Euclid* mission is expected to perform two main surveys, photometric and spectroscopic.

The photometric part should measure photo-z and ellipticities in the optical and near-IR bands (one broad visual band $R + I + Z$ and Y, J, H IR bands), up to ∼24.5 mag in the visual and 24 mag in the IR. The requirement specifications are described in Table 1. The expected number of observed galaxies is of the order of a billion. For this survey we will also use the approximated redshift distribution by Smail et al. (1994), dividing the sample into 12 redshift bins, whose distribution is again convolved with the expected photometric errors, as shown in Fig. 2. The photometric galaxies are distributed in the redshift range $0 < z < 2.5$.

The spectroscopic survey is expected to use a slitless spectrometer which will mainly detect the Hα emission line of galaxies. The spectrometer will have a resolution $\lambda/\Delta\lambda = 500$, giving a redshift uncertainty of $\sigma_z(z) = 0.001(1+z)$. The wavelength range of this instrument will be limited to $1000 < \lambda < 2000$ nm, meaning that only galaxies at $0.5 < z < 2$ will have measurable Hα lines and thus redshifts. The limiting flux is placed at 4×10^{-16} erg s^{-1} cm^{-2}, which combined with the expected success rate of the spectrometer $e = 35$ per cent yields ∼60 million galaxies, using the predicted tabulated calculations by Geach et al. (2010), which was based on the empirical data of the luminosity function of Hα emitter galaxies out to $z = 2$. We will use this tabulated prediction as our fiducial redshift distribution of the sources, consistently with the *Euclid* Study Report specifications (Laureijs 2009). The remaining specifications are again to be found in Table 1. We will finally split the distribution into 12 equally populated redshift bins, as we can see in Fig. 2. Again, we will take the fiducial value of the galactic bias to follow the law $b(z) = \sqrt{1+z}$ following Rassat et al. (2008), which is a good approximation to recent studies from semi-analytic

[1] http://www.darkenergysurvey.org/

[2] http://sci.esa.int/euclid

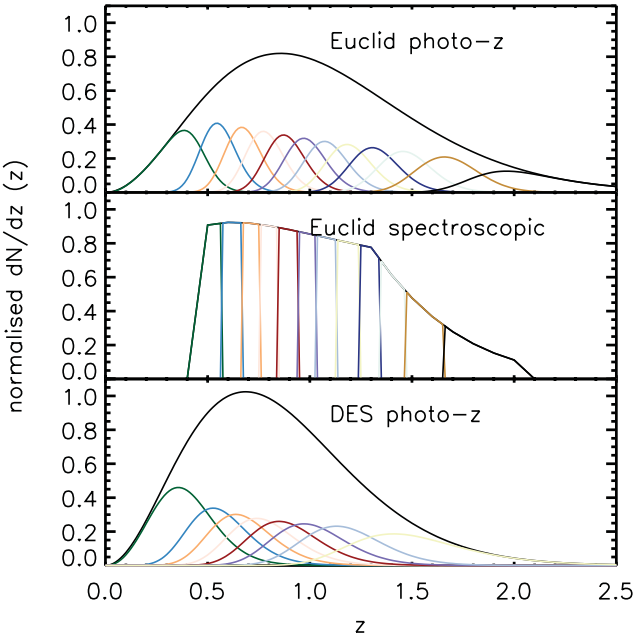

Figure 2. Redshift distributions used for the forecasts of both photometric and spectroscopic data sets for the *Euclid* satellite and the DES. The photo-z distributions are given by an analytic function (Smail et al. 1994), while the spectroscopic part is numerically estimated by Geach et al. (2010), where the *Euclid* specified flux cut is used, 4×10^{-16} erg s^{-1} cm^{-2}. The distributions have been already convolved with the probability density function of redshift measurement errors.

models of galaxy formation such as e.g. by Orsi et al. (2010), but we will study the effect of changing this choice in Section 6.4.

5 FISHER MATRIX FORECASTS

5.1 Formalism

Let us assume, following e.g. Tegmark, Taylor & Heavens (1997) and Tegmark (1997), that we have a data vector $\boldsymbol{x}$, whose elements are random variables which depend on the cosmological parameters $\boldsymbol{\vartheta}$, with a likelihood function $L(\boldsymbol{x}, \boldsymbol{\vartheta})$. Then, if we fix $\mathsf{L} \equiv -\ln L$, the Fisher matrix can be defined as the ensemble average

$$\mathsf{F}_{ij} \equiv \left\langle \frac{\partial^2 \mathsf{L}}{\partial \boldsymbol{\vartheta}_i \, \partial \boldsymbol{\vartheta}_j} \right\rangle, \tag{19}$$

and the Cramér–Rao inequality proves that its inverse represents the best possible covariance matrix for the measurement of the parameters $\boldsymbol{\vartheta}$ with the help of unbiased estimators. In the limit of large data sets, if the distribution becomes Gaussian, this inequality becomes an equality, and the standard deviation of a given parameter $\boldsymbol{\vartheta}_i$ is given by $\sigma(\boldsymbol{\vartheta}_i) = \sqrt{(\mathsf{F}^{-1})_{ii}}$. Under the assumption of a Gaussian likelihood function L, for data of mean $\boldsymbol{\mu} \equiv \langle \boldsymbol{x} \rangle$ and covariance matrix $\boldsymbol{\Sigma} \equiv \langle \boldsymbol{x}\boldsymbol{x}^{\mathrm{T}} \rangle - \boldsymbol{\mu}\boldsymbol{\mu}^{\mathrm{T}}$, the Fisher matrix can be written as the sum of two pieces:

$$\mathsf{F}_{ij} = \frac{1}{2} \mathrm{Tr}\left[\boldsymbol{\Sigma}^{-1} \, \boldsymbol{\Sigma}_{,i} \, \boldsymbol{\Sigma}^{-1} \, \boldsymbol{\Sigma}_{,j} \right] + \boldsymbol{\mu}_{,i}^{T} \, \boldsymbol{\Sigma}^{-1} \, \boldsymbol{\mu}_{,j}, \tag{20}$$

where the commas denote derivatives with respect to the parameters.

To calculate the Fisher matrix, we first need to define our parameter set and its fiducial model, which we take in our base analysis to be for the galaxy clustering case: $\boldsymbol{\vartheta}_0 = \{\Omega_\Lambda = 0.721, \, \Omega_\mathrm{b} = 0.0462, \, \Omega_\mathrm{m} = 0.279, \, h = 0.701, \, n_\mathrm{s} = 0.96, \, \sigma_8 = 0.817, \, w_0 = -1.0, \, w_\mathrm{a} = 0, \, f_\mathrm{NL} = 0, \, b = b_\mathrm{fid}\}$***, where the usual set of cosmological parameters is chosen to match the *WMAP*5 set used in PPH08, the dark-energy equation of state is parametrized as $w(z) = w_0 + w_\mathrm{a} \, z/(1+z)$ (Chevallier & Polarski 2001) and the array b of fiducial value b_fid is a set of nuisance parameters to describe the total scale-independent part of the galaxy bias in each redshift bin (see discussion in Section 2.2). Note that the degeneracy between σ_8 and b is broken by the non-linear corrections to the power spectrum at $k > 0.1\,h$ Mpc^{-1}. For the lensing case, as there is no biasing, we do not include bias nuisance parameters. We will also consider different parameter arrays and extra nuisance parameters in some cases, as described in the following sections.

5.2 Observables

Starting from a 3D random field $\varphi(\boldsymbol{x})$, its 3D power spectrum $P(k)$ can be defined as in equation (2). Similarly, if we consider a 2D random field on a sphere $\xi(\hat{\boldsymbol{n}})$, the angular power spectrum C_l is given by $\langle a_{lm} \, a^*_{l'm'} \rangle = \delta_{ll'} \delta_{mm'} \, C_l$, where a_{lm} are the spherical harmonic coefficients of the field. We will use in our analysis the information contained in all the relevant two-point correlation functions (or power spectra) which are measurable by the DES and *Euclid* missions: we shall consider the 2D power spectra of weak lensing (using the photometric redshift catalogue) and galaxy clustering (using both the photometric and spectroscopic catalogues), and the 3D power spectrum for clustering (spectroscopic), as described below. We will first present the forecasts for each single probe, and finally combine lensing and clustering, including all the relevant two-point correlations and covariances. For all numerical calculations we use the ICOSMO package by Refregier et al. (2011).

An important point is whether this basic Fisher setting, which assumes a Gaussian distribution of the covariance, is accurate enough. More extended calculations have shown that extra terms arise due to the connected matter trispectrum and also to the coupling of large-scale modes which are outside the volume of the survey (Sato et al. 2009; Takada & Jain 2009). The effect of these terms is to reduce the signal-to-noise ratio obtained from the small scales, thus mostly affecting the cosmic shear results. The worsening of the lensing marginalized errors is expected in most cases not to exceed the ∼10 per cent level (Takada & Jain 2009), and thus, while it will be important for data analysis, we will ignore it at the forecasting stage.

5.3 Cosmic shear

We can describe a tomographic galaxy survey of N galaxies in M redshift bins with a set of distribution functions for the galaxy density fluctuations in each ith bin $\mathrm{d}N_i/\mathrm{d}z(z)$, so that the total number of galaxies in the bin is

$$N_i = \int_0^\infty \mathrm{d}z' \, \frac{\mathrm{d}N_i}{\mathrm{d}z'}(z'), \qquad N = \sum_{i=1}^{M} N_i. \tag{21}$$

In a Friedmann–Robertson–Walker universe of total energy density Ω_0 we can write the curvature of the hypersurfaces at constant cosmic time as

$$K = \frac{H_0^2}{c^2}(\Omega_0 - 1), \tag{22}$$

and the comoving angular diameter distance is

$$r_K(r) = \begin{cases} K^{-1/2} \sin(K^{1/2} r) \text{ if } K > 0 \\ r \text{ if } K = 0 \\ (-K)^{-1/2} \sinh(-K^{1/2} r) \text{ if } K < 0, \end{cases} \tag{23}$$

which coincides with the comoving distance r in the flat case. The spatial part of scalar metric perturbations can be decomposed in terms of the solutions of the Helmholtz differential equation (Abbott & Schaefer 1986)

$$(D^2 + k^2)Q(\boldsymbol{r}) = 0\,, \tag{24}$$

where D^2 is the covariant Laplacian in curved space. The eigenfunctions $Q(r)$ generalize the concept of plane waves to curved space–time. Introducing spherical harmonics leads to variable separation and the radial part of the eigenfunctions can be written as $R_{k,l}(r) = \Phi^l_\beta(r)$ (Abbott & Schaefer 1986), $\beta = \sqrt{k^2 + K}$, and $\Phi^\beta_l(r)$ are the ultra-spherical Bessel functions [note that Abbott & Schaefer (1986) write their explicit form in terms of a different radial coordinate].

We can use the weighting kernels for cosmic shear (Bartelmann & Schneider 2001)

$$W^{\epsilon_i}(z) = \frac{1}{N_i}\,\frac{3H_0^2\Omega_{\rm m}}{2\,c^2\,a(z)}\int_z^\infty {\rm d}z'\,\frac{{\rm d}N_i}{{\rm d}z'}(z')\,\frac{r_K[r(z')-r(z)]}{r_K[r(z')]} \tag{25}$$

to calculate the projected 2D cosmic shear power spectrum between the bins i, j as

$$\begin{aligned} C_l^{\epsilon_i\epsilon_j} &= \frac{2}{\pi}\int_0^\infty {\rm d}\beta\,\beta^2\int_0^\infty {\rm d}r_1\,W^{\epsilon_i}(r_1)\,\Phi^\beta_l(r_1) \\ &\times \int_0^\infty {\rm d}r_2\,W^{\epsilon_j}(r_2)\,\Phi^\beta_l(r_2)\;P_{\rm m}(\beta, r_1, r_2)\,, \end{aligned} \tag{26}$$

where $P_{\rm m}$ denotes the cross-spectrum between the matter distribution at two different cosmic epochs corresponding, on our past light cone, to distances r_1 and r_2.[3] This can then be written in the Limber approximation as

$$C_l^{\epsilon_i\epsilon_j} \simeq \int_0^\infty \frac{{\rm d}r}{r_K^2(r)}\,W^{\epsilon_i}[z(r)]\,W^{\epsilon_j}[z(r)]\,P_{\rm m}\left[k = \frac{l+\frac{1}{2}}{r_K(r)}; z(r)\right], \tag{27}$$

and the full 3D matter power spectrum can be calculated e.g. using the halo model formalism of Section 3.

The observed lensing power spectra $\tilde{\mathbf{C}}_l^{\epsilon_i\epsilon_j}$ may be modelled as a sum of the theoretical spectra and the noise. For each pair of redshift bins i, j, we have

$$\tilde{\mathbf{C}}_l^{\epsilon_i\epsilon_j} = C_l^{\epsilon_i\epsilon_j} + N_l^{\epsilon_i\epsilon_j}\,, \tag{28}$$

where the noise term $N_l^{\epsilon_i\epsilon_j}$ is due to the shape noise

$$N_l^{\epsilon_i\epsilon_j} = \delta_{ij}\,\frac{\gamma^2}{\bar{n}_i}\,. \tag{29}$$

Here $\bar{n}_i$ is the angular number density in the ith bin, and γ is the rms shear arising from the intrinsic ellipticities of the galaxies, which is a specification of the survey (see Table 1).

Another important source or errors in cosmic shear are intrinsic alignments: we do expect that a fraction of neighbouring galaxies will have correlated ellipticities due to tidal fields. This effect introduces a systematic into the cosmic shear analysis which has to be taken into account when constraining cosmology from real data (Joachimi & Bridle 2010; Kirk, Bridle & Schneider 2010), while we will neglect it for our forecast as the biasing introduced in the parameter estimation typically does not exceed the ∼10 per cent level.

[3] For a universe with positive curvature the integral over the corrected wavenumber β should be replaced by a discrete sum such that $\beta \geq 3$ and $\beta > l$.

We can now calculate the Fisher matrix element from equation (20) for the case of weak lensing. A common assumption in this case is to identify the observables $\boldsymbol{x}$ with the $a_{lm}^{\epsilon_i\epsilon_j}$ instead of the $C_l^{\epsilon_i\epsilon_j}$, because the a_{lm} are at least in the linear regime Gaussian. This brings in the advantage of cancelling the second term of equation (20), since $\boldsymbol{\mu_i} = \langle a^i_{lm}\rangle = 0$, $\forall\,\{i, l, m\}$. By definition, the variance of each of the $(2l+1)$ $a_{lm}^{\epsilon_i\epsilon_j}$ is given by $\boldsymbol{\Sigma}_{ab} = C_a^{\epsilon_i\epsilon_j}\delta_{ab}$, and thus equation (20) yields (Hu & Jain 2004; Amara & Refregier 2007)

$$F^\epsilon_{\alpha\beta} = f_{\rm sky}\sum_{l=l_{\rm min}}^{l_{\rm max}}\frac{(2l+1)}{2}\,{\rm Tr}\left[\boldsymbol{D}^\epsilon_{l\alpha}\,\left(\tilde{\mathbf{C}}^\epsilon_l\right)^{-1}\,\boldsymbol{D}^\epsilon_{l\beta}\,\left(\tilde{\mathbf{C}}^\epsilon_l\right)^{-1}\right], \tag{30}$$

where $f_{\rm sky}$ is the sky coverage of the survey. Here $\tilde{\mathbf{C}}^\epsilon_l$ is a matrix of dimensions $M \times M$, whose elements are the lensing power spectra at a fixed l between each pair of redshift bins $\tilde{\mathbf{C}}_l^{\epsilon_i\epsilon_j}$. Finally, $\mathbf{D}^\epsilon_{l\alpha}$ is a matrix containing the derivatives of the spectra with respect to each cosmological parameter $\boldsymbol{\vartheta}_\alpha$, whose elements are defined as

$$\mathbf{D}_{l\alpha}^{\epsilon_i\epsilon_j} = \frac{\partial\mathbf{C}_l^{\epsilon_i\epsilon_j}}{\partial\boldsymbol{\vartheta}_\alpha}\,. \tag{31}$$

We shall calculate the forecasted weak-lensing power spectra accordingly with the proposed specifications of the DES and *Euclid* missions, as described above.

The range of multipoles used in the sum of equation (30) [$l_{\rm min}$, $l_{\rm max}$] is a sensitive issue, since we would like to set it as broad as possible to use all available information. The value of $l_{\rm max}$ is related to how far we can trust our modelling of the non-linear regime. In the case of weak lensing, as we are free from the problems of non-linear biasing, we take $l_{\rm max} = 20\,000$, which gives $k^\epsilon_{\rm max} \simeq 8.5\,h\,{\rm Mpc}^{-1}$ at $z = 1$ assuming the best-fitting *WMAP* cosmology. Again this assumes that the effect of baryons on the matter power spectrum can be either ignored or accurately modelled. Another issue which may affect the analysis at these small scales is the non-Gaussian contribution to the covariance; we do not account for these corrections whose effects are expected to be small, as discussed at the end of Section 5.2. Further, the theory calculation in this regime is strongly dependent on the parameters of the halo model, such as the concentration, the subhalo distribution, etc. We will discuss in Section 6.7 the effects of changing this limit, where we will find that with more conservative choices the differences are still small (see Fig. 7): for example, the marginalized error on $f_{\rm NL}$ including *Planck* priors degrades in this case by only 15 per cent between $l_{\rm max} = 20\,000$ and $l_{\rm max} = 5000$.

The minimum multipole used is also important. It is expected that the accuracy of the Limber approximation will deteriorate for $l \to 0$. Here, to be complete and conservative, we will only use this approximation at small scales, for $l \geq 200$, and will use the complete exact formula at larger scales (using the linear approximation for the cross-spectrum at two times). In this way, we can extend the calculation all the way to the largest scales, and will take $l_{\rm min} = 5$. However, as shown in Fig. 3, the exact calculation is important only for the case of galaxy clustering, where it cannot be neglected as in the presence of non-Gaussianity, the large scales (small multipoles) are the most affected by the scale-dependent bias. Note that the Limber approximation overestimates the power on the largest scales for positive $f_{\rm NL}$ while it underestimates it for a Gaussian model. Since the Fisher matrix is computed taking the derivatives of the signal with respect to model parameters, the Limber approximation is rather inaccurate for a fiducial with $f_{\rm NL} = 0$.

We can see the resulting 2D marginalized forecasts for a *Euclid*-like survey for local PNG in Fig. 4, and all the marginalized 1D

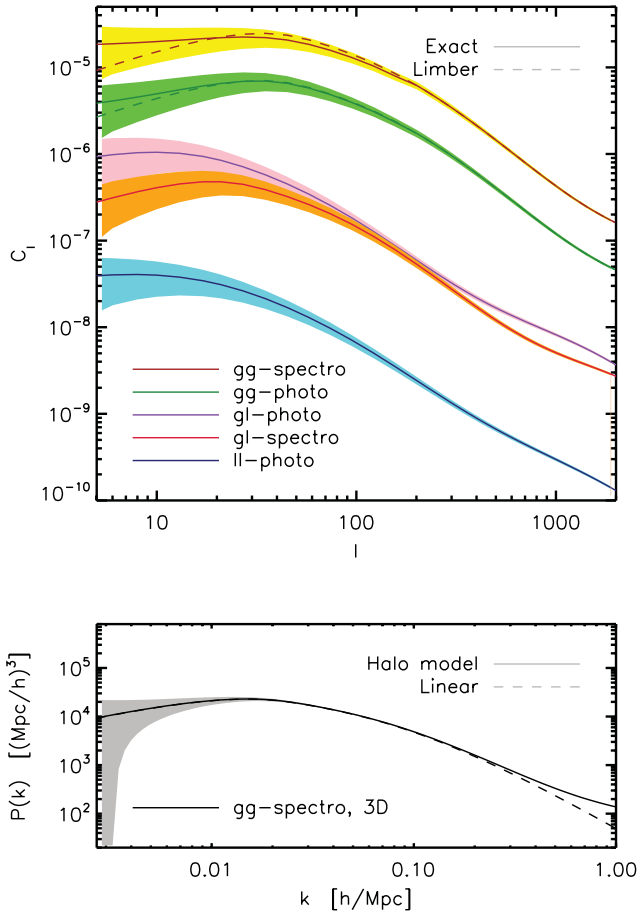

Figure 3. Summary of all observables used in this work for the fiducial model described in Section 5. In the upper panel, we show the projected 2D spectra for lensing, galaxy clustering and their cross-spectra for photometric and spectroscopic surveys. In the bottom panel we show the 3D galaxy power spectrum. Note that the Limber approximation is inaccurate on large scales for the galaxy–galaxy spectra; on the other hand, the linear-theory power spectrum is inaccurate on the small scales (dashed lines). For the lensing and galaxy-lensing cases, the Limber approximation works well due to the wider distribution of the sources. The shaded areas represent cosmic-variance errors for a half-sky survey. A redshift bin centred around $z = 1$ is used in all cases.

error bars in Fig. 6. The results are also summarized in detail in Tables 2, 3 and 4 for the three PNG configurations. We present the results for the DES in Table 5 for the local case.

5.4 Galaxy clustering, 3D

The second observable we consider is the clustering of galaxies, which is typically measured via the 3D power spectrum using the spectroscopic part of the survey. In this case we then consider measurements of the galaxy power spectrum for each ith redshift bin $P_{g_i}(k) = b^2_{g_i}(k)\, P_m(k)$, where $b_{g_i}(k) = \left[b_{g_i} + \Delta b_i(k)\right]$ is the galaxy bias of the ith bin. The scale-dependent part of the bias is computed as explained in Section 2, while we consider the scale-independent part as a nuisance parameter to be marginalized over. Due to the high accuracy of the spectroscopic redshifts, we can ignore any cross-spectra between different redshift bins, since the redshift distribution functions will have negligible overlap, and we will also neglect the effect of covariances due to the longitudinal modes. We also include corrections due to two effects, as described e.g. by Seo & Eisenstein (2003) and Song & Percival (2009): the redshift-space distortions and the Alcock–Paczynski effect.

Redshift-space distortions are caused by the peculiar motion of galaxies (Kaiser 1987) and produce an enhancement of the overdensity field δ proportional to $(1 + f\mu^2)$, where f is the logarithmic derivative of the growth function with respect to the expansion factor a, and μ is the cosine of the angle to the line of sight. If we exclude velocity bias, the correction of the power spectrum becomes (see e.g. Peacock & Dodds 1994)

$$P^s_{g_i}(\boldsymbol{k}) = P_{g_i}(\boldsymbol{k}) \left[1 + \mu^2 \beta_{g_i}(k)\right]^2 F\left(\frac{k^2\mu^2\sigma^2_{v,i}}{H^2(z)}\right), \tag{32}$$

where $\beta_{g_i}(k) = f/b_{g_i}(k)$, $\sigma_{v,i}$ is the 1D pairwise velocity dispersion of the galaxies and $H(z)$ is the Hubble function. We will use for the function F the form $F(x) = e^{-x}$ (Percival et al. 2007). This gives a small correction whose cosmological content can be ignored and treated as a nuisance parameter to be marginalized over (Song & Percival 2009). In more detail, we can decompose the velocity dispersion in each redshift bin i as the sum in quadrature of an intrinsic term coming from the finger-of-god (FoG) effect, $\sigma_{FoG,i}$, and a term due to the redshift uncertainty of the survey σ_z given in Table 1, as

$$\sigma^2_{v,i} = (1+z)^2 \left[\frac{\sigma^2_{FoG,i}}{2} + c^2\sigma^2_z\right]. \tag{33}$$

Since the redshift evolution of this intrinsic term is uncertain, it would be desirable to introduce independent parameters $\sigma_{FoG,i}$ for each redshift bin and then marginalize over them. However, we found that, given that this factor enters the calculation together with the function $1/H(z)$ in equation (32), leaving these parameters free to change brings in severe degeneracies with any other cosmological parameter which is able to alter $H(z)$. For this reason, we have been forced to adopt a less conservative approach, where we choose a functional form

$$\sigma_{FoG}(z) = \sigma_{FoG,0}\sqrt{1+z} \tag{34}$$

and only marginalize on one parameter $\sigma_{FoG,0}$. This scaling with redshift assumes that the observed galaxies reside in haloes with nearly constant mass. It is also possible to make this approach more conservative, e.g., by considering more complicated parametrizations for this function, but again the constraints on all parameters which enter the Hubble expansion $H(z)$ will worsen dramatically. As a fiducial value, we take $\sigma_{FoG,0} = 250$ km s^{-1}, as the *Euclid* spectroscopic sample is made of star-forming galaxies, which are not generally located in massive haloes, and are thus expected to have a low-velocity dispersion.

On the other hand, the Alcock–Paczynski effect occurs since to infer galaxy distances from the observed redshifts and positions we have to use a reference cosmology, which is different from the 'true' one. This can be corrected, and the true values are (Seo & Eisenstein 2003)

$$\begin{aligned} k_{true} &= \left[k^2_{ref}\left(1-\mu^2_{ref}\right)\frac{D^2_{ref}(z)}{D^2_{true}(z)} + (k_{ref}\,\mu_{ref})^2\frac{H^2_{true}(z)}{H^2_{ref}(z)}\right]^{1/2} \\ \mu_{true} &= k_{ref}\,\mu_{ref}\,\frac{H^2_{true}(z)}{H^2_{ref}(z)}\,\frac{1}{k_{true}}. \end{aligned} \tag{35}$$

Here we will identify the reference cosmology with our fiducial model for simplicity. Finally, the observable corrected power spectrum, which we will use to calculate the Fisher matrix, is given

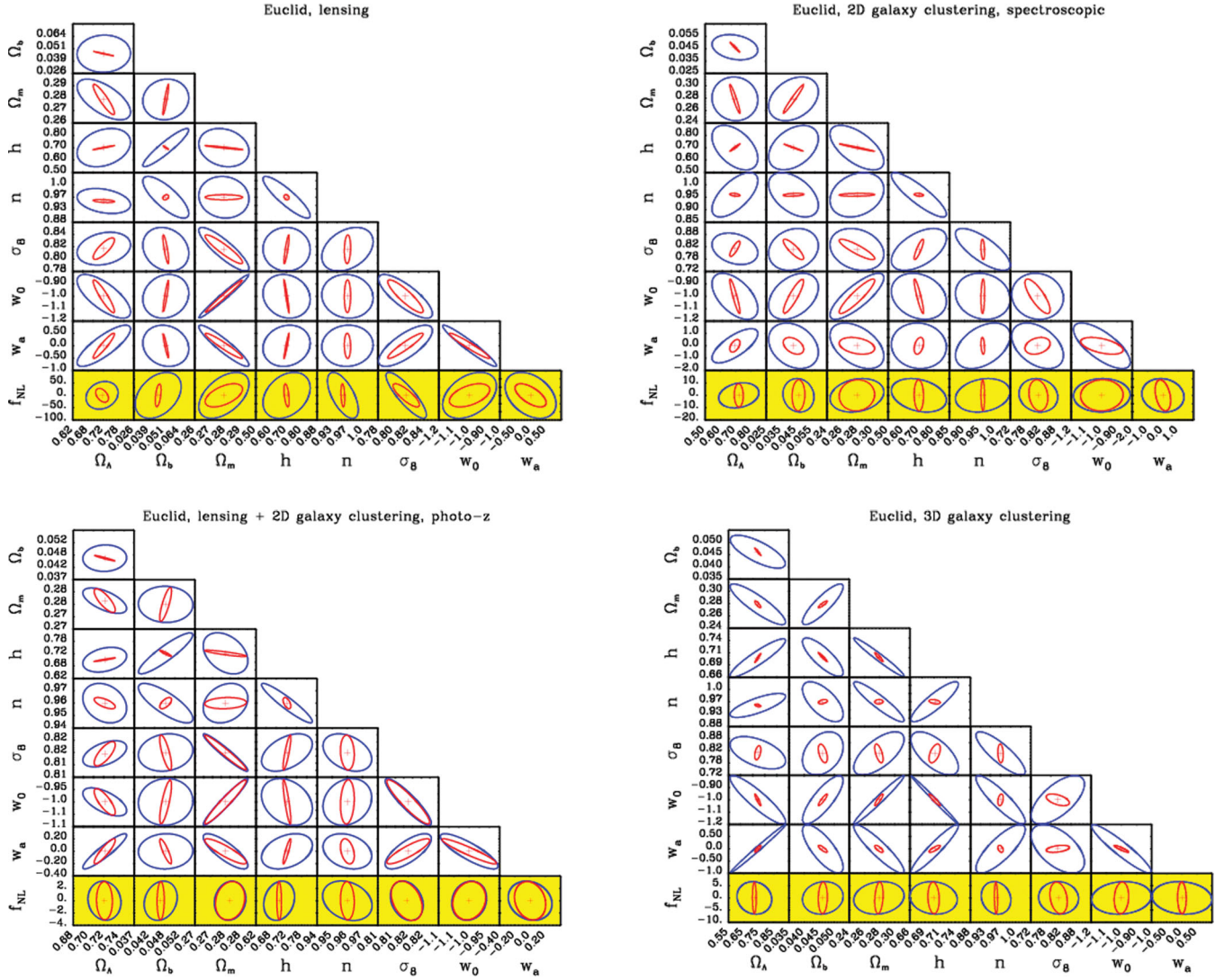

Figure 4. Fisher matrix forecasts for the *Euclid*-like survey, using weak lensing (photometric survey), 2D galaxy clustering (photometric and spectroscopic surveys) and 3D galaxy clustering (spectroscopic only). For lensing, the used multipoles are from $l_{\rm min} = 5$ to $l_{\rm max} = 20\,000$, while for clustering the maximum mode is $k_{\rm max} = 0.15\,h\,{\rm Mpc}^{-1}$ at $z = 0$. The forecasted posteriors are marginalized over the other not shown parameters. The blue ellipses refer to *Euclid* only, while in red we show the results including *Planck* priors. Note that the axes' ranges are different, which is necessary given the different constraining power of the different observables.

by

$$\tilde{P}_{{\rm g}_i}(k_{\rm true},\mu_{\rm true}) = \frac{D^2_{\rm true}(z)\,H_{\rm ref}(z)}{D^2_{\rm ref}(z)\,H_{\rm true}(z)}\,b^2_{{\rm g}_i}(k_{\rm true})\left[1+\beta_{{\rm g}_i}(k_{\rm true})\,\mu^2_{\rm true}\right]^2 \times P_{\rm m}(k_{\rm true})\,F\left(\frac{k^2_{\rm true}\mu^2_{\rm true}\sigma^2_{v,i}}{H^2(z)}\right) + P_{{\rm shot},{\rm g}_i}\,. \quad (36)$$

The last term on the right-hand side is due to shot noise. If this is Poissonian, its fiducial value is $P_{{\rm shot},{\rm g}_i} = 1/\bar{n}_i$. However, it is possible to be more conservative and assume that the mean galaxy density is not perfectly known, or that there are other sources of shot noise. We will include these unknown contributions by introducing one additional nuisance parameter for it in each bin, and marginalizing over them.

The Fisher matrix calculation is here based on the works by Feldman, Kaiser & Peacock (1994) and Tegmark (1997) amongst others. In the 3D case, it is customary to identify each of the observables $\boldsymbol{x}_i$ with the average power in a thin shell of radius k_i in Fourier space, of width ${\rm d}k_i$ and volume $V_i = 4\pi k_i^2 {\rm d}k_i/(2\pi)^3$. In this case, for each redshift bin a and angle μ_b we have non-zero means $\boldsymbol{\mu}_i \simeq \tilde{P}_{{\rm g}_a}(k_i,\mu_b)$ and covariances

$$\boldsymbol{\Sigma}_{ij}(\mu_b) \simeq 2\frac{\tilde{P}_{{\rm g}_a}(k_i,\mu_b)\tilde{P}_{{\rm g}_a}(k_j,\mu_b)}{V_i\,V_{\rm eff}(k_i,\mu_b)}\delta_{ij}, \quad (37)$$

where the effective volume is

$$V_{\rm eff}(k,\mu) = \int\left[\frac{\bar{n}(\boldsymbol{r})\tilde{P}_{{\rm g}_a}(k,\mu)}{1+\bar{n}(\boldsymbol{r})\tilde{P}_{{\rm g}_a}(k,\mu)}\right]^2 {\rm d}^3r, \quad (38)$$

and $n(\boldsymbol{r})$ is the selection function of the survey. The dominant term of the Fisher matrix is now the second one in equation (20) which, adding the effects of redshift-space distortions, can be finally written as (Seo & Eisenstein 2003)

$$\mathbf{F}^{\rm 3D}_{\alpha\beta} = \int_{-1}^{1}\int_{k_{\rm min}}^{k_{\rm max}} \frac{\partial\ln\tilde{P}(k,\mu)}{\partial\vartheta_\alpha}\frac{\partial\ln\tilde{P}(k,\mu)}{\partial\vartheta_\beta}\,V_{\rm eff}(k,\mu)\,\frac{\pi k^2}{(2\pi)^3}\,{\rm d}k\,{\rm d}\mu. \quad (39)$$

Since here we are not accounting for bias non-linearities nor scale dependences at small scales, our calculated galaxy power spec-

Table 2. Expected marginalized errors from *Euclid*-like data for all parameters for the local non-Gaussianity case. The numbers within parentheses include the forecasted priors from the *Planck* CMB temperature power spectrum. The additional line for the 3D clustering includes marginalization over the shot noise parameters in each redshift bin.

Euclid data	$\sigma(\Omega_\Lambda)$	$\sigma(\Omega_b)$	$\sigma(\Omega_m)$	$\sigma(h)$	$\sigma(n)$	$\sigma(\sigma_8)$	$\sigma(w_0)$	$\sigma(w_a)$	$\sigma(f_{NL})$
Lensing	0.035 (0.014)	0.012 (0.0016)	0.011 (0.0081)	0.10 (0.011)	0.033 (0.0040)	0.017 (0.011)	0.12 (0.087)	0.55 (0.34)	37 (19)
2D clustering photometric	0.064 (0.037)	0.0085 (0.0043)	0.050 (0.026)	0.12 (0.033)	0.027 (0.0036)	0.042 (0.028)	0.51 (0.29)	1.6 (0.94)	3.2 (3.0)
2D clustering spectroscopic	0.087 (0.019)	0.0067 (0.0027)	0.023 (0.015)	0.12 (0.020)	0.059 (0.0039)	0.042 (0.020)	0.12 (0.092)	0.91 (0.33)	6.7 (6.2)
3D clustering	0.063 (0.0079)	0.0041 (0.000 96)	0.018 (0.0036)	0.020 (0.006)	0.024 (0.0033)	0.030 (0.016)	0.10 (0.030)	0.53 (0.090)	4.4 (4.2)
– +shot noise	0.076 (0.0082)	0.0044 (0.000 98)	0.022 (0.0036)	0.026 (0.0062)	0.032 (0.0035)	0.044 (0.019)	0.13 (0.031)	0.65 (0.094)	4.4 (4.2)
Lensing + 2D photometric	0.014 (0.0069)	0.0035 (0.000 77)	0.0034 (0.0032)	0.033 (0.0050)	0.011 (0.0030)	0.0035 (0.0032)	0.037 (0.035)	0.19 (0.13)	2.8 (2.6)
Lensing + 2D spectroscopic	0.017 (0.0083)	0.0049 (0.000 94)	0.0048 (0.0044)	0.044 (0.0064)	0.014 (0.0035)	0.0051 (0.0047)	0.052 (0.047)	0.24 (0.17)	4.9 (4.6)

Table 3. Expected marginalized errors from *Euclid*-like data for all parameters, for the orthogonal configuration. The constraints on f_{NL} from the scale-dependent bias are degraded, while the results from lensing are in this case stronger.

Euclid data	$\sigma(\Omega_\Lambda)$	$\sigma(\Omega_b)$	$\sigma(\Omega_m)$	$\sigma(h)$	$\sigma(n)$	$\sigma(\sigma_8)$	$\sigma(w_0)$	$\sigma(w_a)$	$\sigma(f_{NL})$
Lensing	0.035 (0.014)	0.013 (0.0016)	0.011 (0.0080)	0.10 (0.011)	0.035 (0.0040)	0.017 (0.011)	0.12 (0.087)	0.55 (0.33)	4.9 (2.4)
2D clustering photometric	0.064 (0.039)	0.010 (0.0046)	0.051 (0.028)	0.12 (0.035)	0.040 (0.0039)	0.11 (0.050)	0.50 (0.31)	1.6 (0.96)	28 (8.1)
2D clustering spectroscopic	0.097 (0.020)	0.010 (0.0029)	0.023 (0.016)	0.15 (0.021)	0.086 (0.004)	0.17 (0.045)	0.12 (0.096)	0.95 (0.32)	50 (10)
3D clustering + shot	0.076 (0.0083)	0.0045 (0.0010)	0.022 (0.0036)	0.026 (0.0062)	0.032 (0.0036)	0.23 (0.043)	0.13 (0.031)	0.65 (0.095)	56 (12)
Lensing + 2D photometric	0.014 (0.0076)	0.0040 (0.000 88)	0.0048 (0.0039)	0.034 (0.0058)	0.012 (0.0033)	0.0096 (0.0064)	0.049 (0.041)	0.24 (0.19)	2.9 (1.6)
Lensing + 2D spectroscopic	0.018 (0.0087)	0.0053 (0.000 99)	0.0053 (0.0048)	0.047 (0.0068)	0.018 (0.0036)	0.010 (0.0071)	0.055 (0.050)	0.26 (0.20)	3.4 (1.9)

Table 4. Expected marginalized errors from *Euclid*-like data for all parameters for the equilateral configuration. The bias is kept scale- independent here. The constraints on $f_{\rm NL}$ from galaxy clustering are strongly weakened, as the bias is now scale independent. The results from lensing remain instead strong.

Euclid data	$\sigma(\Omega_\Lambda)$	$\sigma(\Omega_{\rm b})$	$\sigma(\Omega_{\rm m})$	$\sigma(h)$	$\sigma(n)$	$\sigma(\sigma_8)$	$\sigma(w_0)$	$\sigma(w_{\rm a})$	$\sigma(f_{\rm NL})$
Lensing	0.035 (0.014)	0.012 (0.0016)	0.011 (0.0082)	0.10 (0.011)	0.031 (0.0040)	0.017 (0.012)	0.12 (0.088)	0.54 (0.34)	17 (9.2)
2D clustering photometric	0.085 (0.041)	0.0098 (0.0047)	0.057 (0.029)	0.12 (0.036)	0.038 (0.0039)	0.12 (0.057)	0.51 (0.32)	1.7 (1.0)	100 (30)
2D clustering spectroscopic	0.11 (0.020)	0.0069 (0.0028)	0.024 (0.016)	0.12 (0.020)	0.073 (0.0040)	0.16 (0.048)	0.14 (0.094)	1.1 (0.31)	150 (35)
3D clustering + shot	0.079 (0.0083)	0.0045 (0.000 99)	0.024 (0.0036)	0.027 (0.0062)	0.038 (0.0036)	0.25 (0.043)	0.14 (0.031)	0.69 (0.094)	220 (37)
Lensing + 2D photometric	0.014 (0.0076)	0.0040 (0.000 88)	0.0052 (0.0039)	0.033 (0.0058)	0.013 (0.0033)	0.011 (0.0066)	0.050 (0.041)	0.24 (0.18)	12 (6.5)
Lensing + 2D spectroscopic	0.017 (0.0090)	0.0050 (0.0010)	0.0058 (0.0049)	0.044 (0.0071)	0.015 (0.0035)	0.011 (0.0076)	0.057 (0.051)	0.28 (0.21)	11 (7.0)

Table 5. Expected marginalized errors for all parameters for the DES for the local PNG case. The numbers within parentheses include the forecasted priors from the *Planck* CMB temperature power spectrum.

DES data	$\sigma(\Omega_\Lambda)$	$\sigma(\Omega_{\rm b})$	$\sigma(\Omega_{\rm m})$	$\sigma(h)$	$\sigma(n)$	$\sigma(\sigma_8)$	$\sigma(w_0)$	$\sigma(w_{\rm a})$	$\sigma(f_{\rm NL})$
Lensing	0.20 (0.033)	0.036 (0.0036)	0.037 (0.022)	0.32 (0.027)	0.13 (0.0042)	0.063 (0.030)	0.41 (0.26)	2.6 (0.96)	150 (46)
2D clustering photometric	0.23 (0.050)	0.019 (0.0056)	0.10 (0.033)	0.34 (0.042)	0.10 (0.0041)	0.058 (0.039)	0.82 (0.44)	3.3 (2.0)	12 (11)
Lensing + 2D photometric	0.062 (0.013)	0.014 (0.0014)	0.0082 (0.0074)	0.12 (0.010)	0.039 (0.0041)	0.0094 (0.0086)	0.093 (0.090)	0.61 (0.35)	8.6 (8.2)
Simple sum	0.12 (0.028)	0.014 (0.0030)	0.024 (0.018)	0.13 (0.022)	0.040 (0.0041)	0.023 (0.018)	0.28 (0.22)	1.6 (0.77)	11 (10)

trum is exact only in the linear regime; therefore to limit the effects of non-linearities, in particular for what concerns galaxy biasing (see Roth & Porciani 2011), we cut our calculation at a scale $k_{\rm max} = 0.15\,h\,{\rm Mpc}^{-1}$ at $z = 0$. We can then make this limit more meaningful by evolving it as a function of z, by imposing the condition

$$\sigma^2(z) = \int_{k_{\rm min}}^{k_{\rm max}(z)} \frac{{\rm d}k}{2\pi^2} k^2 P_0(k, z) = {\rm const.} \tag{40}$$

Here we have set $k_{\rm min} = 10^{-3}\,h\,{\rm Mpc}^{-1}$, as this roughly corresponds to the effective volume of our *Euclid*-like survey. The obtained values of $k_{\rm max}(z)$ can then be translated into $l_{\rm max}$ by using the approximation

$$l^{\epsilon}_{\rm max} \simeq k^{\epsilon}_{\rm max}\, r_K - 1/2\,. \tag{41}$$

In this way the choice $k_{\rm max} = 0.15\,h\,{\rm Mpc}^{-1}$ at $z = 0$ translates into imposing that the variance is fixed to $\sigma^2(z) = 0.36$, i.e. the rms of the perturbations is $\sigma \simeq 0.6$, meaning that perturbation theory is generally respected (peaks of order $\delta \sim 1$ are rare), and so this is a conservative choice. To be even more conservative, we impose that in any case the smallest scale cannot go above $k^{\rm abs}_{\rm max} = 0.3\,h\,{\rm Mpc}^{-1}$, as was similarly done by Wang et al. (2010). In practice, this maximum k is reached around $z \simeq 1$. We have checked that at this redshift, the non-linear contribution to the power spectrum does not exceed ∼13 per cent and is decreasing at higher redshifts.

We can see the results for *Euclid*, local PNG in Fig. 4, and all the marginalized 1D error bars in Fig. 6. The results are also summarized in detail in Tables 2, 3 and 4 for the three PNG configurations. We remind the reader that in this case the results are marginalized over nuisance parameters for the Gaussian part of galaxy bias, the FoG effect and where indicated also shot noise.

5.5 Galaxy clustering, 2D

In addition to the 3D galaxy clustering in redshift space discussed in the previous section, we have also calculated forecasts for the projected clustering on the celestial sphere. This is done for several reasons: first, in this way it is much easier to combine the results with weak lensing in a consistent way, as we will discuss in the next section. Secondly, this allows us to directly compare the performances of the photometric and spectroscopic surveys of the *Euclid*-like mission in terms of galaxy clustering: a priori, it is not obvious whether the tighter constraints on cosmological parameters could come from the more numerous galaxies with photometric redshifts or from the less rich but more accurately located spectroscopic sample. Thirdly, we can thus present forecasts for the galaxy clustering of the DES, which does not include, at least in its initial form, a spectroscopic survey.

The 2D galaxy spectrum between a pair of redshift bins i, j is a projection of the 3D spectrum which, in analogy with the weak-lensing case above, can be written as

$$\begin{aligned} \mathbf{C}_l^{{\rm g}_i{\rm g}_j} = \frac{2}{\pi} \int_0^\infty {\rm d}\beta\, \beta^2 \int_0^\infty {\rm d}r_1\, W^{{\rm g}_i}(\beta, r_1)\, \Phi_l^\beta(r_1) \\ \times \int_0^\infty {\rm d}r_2\, W^{{\rm g}_j}(\beta, r_2)\, \Phi_l^\beta(r_2)\; P_{\rm m}(\beta, r_1, r_2). \end{aligned} \tag{42}$$

This can then be written in the Limber approximation in analogy with equation (27). The sources in a bin i are now given by

$$W^{{\rm g}_i}(\beta, r) = \frac{1}{N_i} \frac{{\rm d}N_i}{{\rm d}r}(r)\, b_{{\rm g}_i}[r, k(\beta)]. \tag{43}$$

The observed 2D galaxy power spectra $\tilde{\mathbf{C}}_l^{{\rm g}_i{\rm g}_j}$ can again be modelled as a sum of the theoretical spectra and the noise. For each pair of redshift bins i, j, we have

$$\tilde{\mathbf{C}}_l^{{\rm g}_i{\rm g}_j} = \mathbf{C}_l^{{\rm g}_i{\rm g}_j} + \mathbf{N}_l^{{\rm g}_i{\rm g}_j}, \tag{44}$$

where the noise term $\mathbf{N}_l^{{\rm g}_i{\rm g}_j}$ is now due to the shot noise:

$$\mathbf{N}_l^{{\rm g}_i{\rm g}_j} = \delta_{ij} \frac{1}{\bar{n}_i}\,. \tag{45}$$

The Fisher matrix in this case is defined exactly as in equation (30) above, with the lensing power spectra replaced by the galaxy–galaxy ones.

For completeness and for the purpose of comparison, we will use in this case both photometric and spectroscopic data sets for *Euclid*, as described in the above sections. Note that, as already mentioned for the 3D case, since here we are not accounting for bias non-linearities nor scale dependences at small scales, equation (42) is exact only in the linear regime; therefore, we use the same limit on the small scales described above for both cases, i.e. we assume now $k_{\rm max} = 0.15\,h\,{\rm Mpc}^{-1}$ at $z = 0$, evolved in redshift as in the 3D case described above. We will discuss in Section 6.7 the effects of changing this limit. Also in this case, we use the Limber approximation for $l \geq 200$, and the exact calculation for larger scales, so that we can extend the analysis all the way to $l_{\rm min} = 5$. See Fig. 3 for a comparison of the Limber power spectra with the exact calculations for the range of projected observables we consider. Here we can see how the Limber approximation departs from the exact calculation at large scales for the galaxy spectra.

We can see the results in Fig. 4 for both photometric and spectroscopic parts of *Euclid* for the local PNG case, and all the marginalized 1D error bars in Fig. 6. The results are also summarized in detail in Tables 2, 3 and 4 for the three PNG configurations. These results are marginalized over nuisance parameters for the Gaussian part of galaxy bias.

5.6 Combined results

Finally, we shall combine the constraints from lensing and galaxy clustering in the projected case, as it is the only case where it is straightforward to do it consistently. Since the two probes are based on the same density field, it is crucial to correctly include their full covariance; on the other hand, it is also possible to include the lensing–clustering cross-correlation as a signal. These two operations can be achieved by considering one single Fisher matrix, as done by Hu & Jain (2004), and defined by

$$\mathbf{F}^x_{\alpha\beta} = f_{\rm sky} \sum_{l=l_{\rm min}}^{l_{\rm max}} \frac{(2l+1)}{2} {\rm Tr}\left[\boldsymbol{D}^x_{l\alpha} \left(\tilde{\mathbf{C}}^x_l\right)^{-1} \boldsymbol{D}^x_{l\beta} \left(\tilde{\mathbf{C}}^x_l\right)^{-1} \right]. \tag{46}$$

In this case, the matrices $\tilde{\mathbf{C}}^x_l$, $\tilde{\boldsymbol{D}}^x_l$ will contain all the combinations of 2D spectra: $\tilde{\mathbf{C}}_l^{\epsilon_i\epsilon_j}$, $\tilde{\mathbf{C}}_l^{\epsilon_i{\rm g}_j}$, $\tilde{\mathbf{C}}_l^{{\rm g}_i{\rm g}_j}$. Note that the noise on the cross-spectra is zero, since we are assuming that shape and Poisson noise are uncorrelated: $\tilde{\mathbf{C}}_l^{\epsilon_i{\rm g}_j} = \mathbf{C}_l^{\epsilon_i{\rm g}_j}$.

We consider in this case the combination of the lensing forecasts using the photometric catalogue with both photometric and spectroscopic galaxy clustering. The combined results for *Euclid* are reported for both cases in Fig. 5 for local PNG and summarized in detail in Tables 2, 3 and 4 for the three PNG configurations. These results are marginalized over nuisance parameters for the Gaussian part of galaxy bias.

All the marginalized 1D error bars are shown in Fig. 6 for the local PNG case. We also present the results obtained for the DES

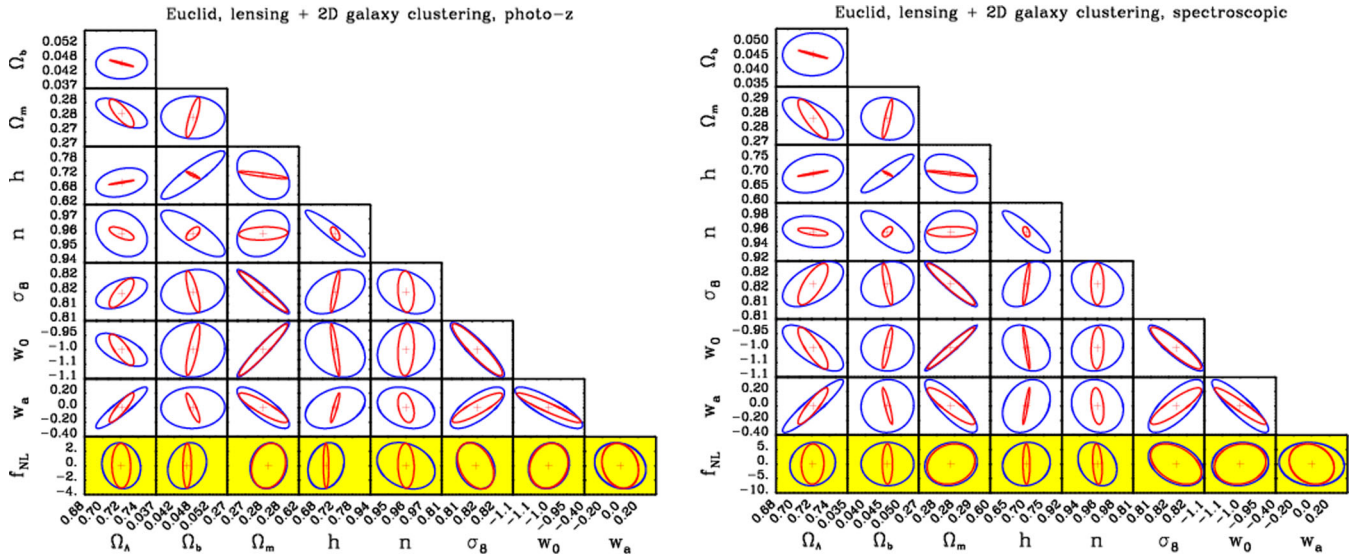

Figure 5. Full combination of *Euclid*-like data. Fisher matrix forecasts for the combination of weak lensing (photometric survey) plus 2D galaxy clustering (photometric and spectroscopic surveys). Both the range of used multipoles and the colour coding of the ellipses are as described above for the previous plots.

specifications in Table 5 for the local case. We can also see from Table 6 the comparison between naively summing the lensing and clustering Fisher matrices versus performing the full analysis. We can see here that the forecasted errors for the complete combination are smaller than the simplistic sum, on average by a factor of 2. This points us to conclude that the effect of including all the cross-spectra between lensing and clustering as observable signals in the analysis is more important than the degrading effect which is produced by considering the covariances between the clustering and lensing autocorrelations.

6 DISCUSSION AND ADDITIONAL ANALYSES

We present here some additional results which extend our basic analysis of the previous section and discuss some important complementary issues.

6.1 Additional priors

We first study the effect of adding as a prior the forecasted constraints for the *Planck* CMB satellite, whose results will be available by the time these surveys are completed. The Fisher matrix for *Planck* is obtained in analogy with the galaxy 2D case, and the CMB spectra are calculated with the CAMBCMB code(Lewis, Challinor & Lasenby 2000). We describe this in more detail in Appendix B.[4]

The reduced likelihood intervals are also shown in Figs 4–6. We can see that in this case all the constraints improve greatly, with the significant exception of $f_{\rm NL}$. This is because the *Planck* prior we are using is only taking into account the CMB power spectrum, and not the bispectrum, which is where most information on non-Gaussianity is. In other words, our *Planck* Fisher matrix does not depend on $f_{\rm NL}$ at all.

6.2 Summary of the results

We summarize in Fig. 6 all the marginalized 1D constraints which are obtained from the Fisher matrix analysis in the Gaussian and non-Gaussian cases, with and without the imposition of the *Planck* priors. Here we can also see a comparison between the forecasted future results for *Euclid* and the current constraints from the *WMAP* satellite and other probes. It is particularly interesting to note that the errors on the standard cosmological parameters remain largely unchanged when $f_{\rm NL}$ is added. We also found by looking at the covariant elements in the Fisher matrices that $f_{\rm NL}$ is almost completely uncorrelated from the other parameters, which is due to the very particular scale-dependent behaviour which it produces on the observables.

6.3 The different shapes of non-Gaussianity

We now focus on the differences between the three types of PNG described in Section 2: namely the local, equilateral and orthogonal bispectrum shapes. In all these cases PNG alters both the matter power spectrum through the halo mass function (which depends on the skewness of the linear density field) and the galaxy bias.

The different forms for the skewness are described in Appendix A, while we use the expressions for the corrections to the bias in equations (9) derived by Schmidt & Kamionkowski (2010) and improved by Desjacques et al. (2011). These expressions are approximated but are rather accurate at small wavenumbers for the local and orthogonal cases (Schmidt & Kamionkowski 2010) where the scale-dependent part of the bias scales as k^{-2} and k^{-1}, respectively. On the other hand, in the equilateral case, the bias becomes asymptotically constant for small k and the only scale dependence of the bias comes from the linear transfer function appearing in the function α. Unfortunately, the numerical factor $\frac{\sigma^2_{R,-2}}{\sigma^2_{R,0}}$ which appears in the bias correction is typically large, so that the scale-dependent deviations sourced by $T(k)$ easily outweigh the effect of the mass-density skewness in the matter power spectrum at large k. This is not ideal, as equations (9) have been derived with the peak-background split technique, which should only hold on large scales. For this reasons we decide here to be conservative, and to keep the bias fully scale independent in the equilateral case, fixing it at all scales to its small-k value.

[4] We acknowledge Jochen Weller for supplying us with the CMB Fisher matrix code.

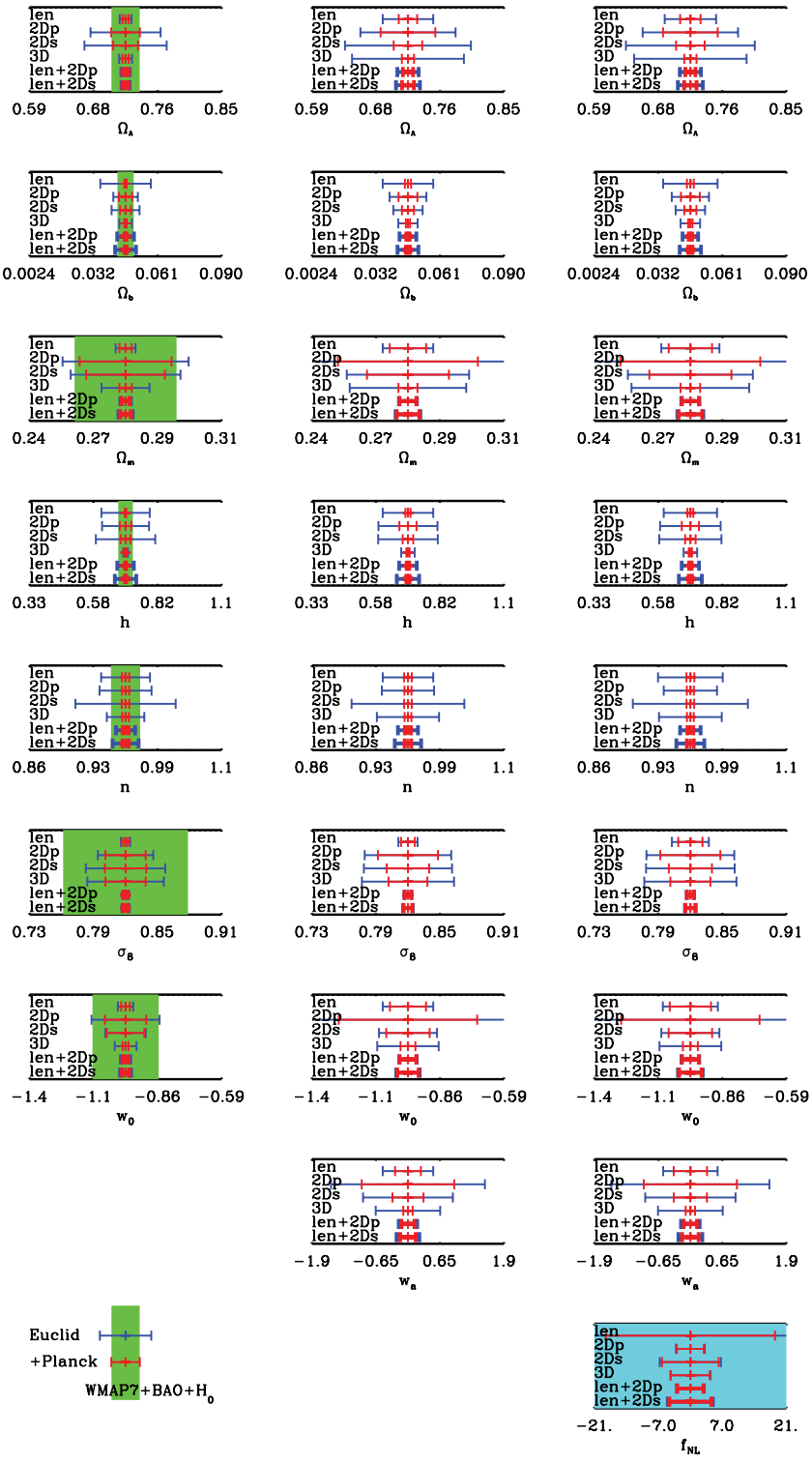

Figure 6. Marginalized 1D errors from *Euclid*-like data on each cosmological parameter, for a wCDM model (left), and with the addition of w_a (centre) and f_{NL} (right), for the local configuration. For each parameter, we show the result obtained using all four observables and their combinations. The error bars in red include the *Planck* priors. We also overlay the current constraints from *WMAP*7, baryon acoustic oscillations and H_0 measurements in light green, and the size of the current uncertainty on f_{NL} from the CMB bispectrum in cyan. The results on the standard parameters remain unaffected by the addition of f_{NL}.

The marginalized constraints on the non-Gaussianity parameter f_{NL} are shown in Table 7 for the three configurations. Here we can see that the constraints in the orthogonal and equilateral cases get generally weaker for the clustering probes due to the reduced importance of the scale-dependent bias. The constraints from lensing remain nevertheless quite strong, as the effect of the skewness is independent from the bias; the results for the combined probes are therefore very promising for all the configurations.

As a further note of caution, we have nevertheless to stress that, differently from the local case, we could not check with *N*-body

Table 6. Comparison of summing two matrices obtained for *Euclid* versus combining them appropriately including all the cross-signals. Results between parentheses include *Planck* priors.

Euclid data	$\sigma(\Omega_\Lambda)$	$\sigma(\Omega_b)$	$\sigma(\Omega_m)$	$\sigma(h)$	$\sigma(n)$	$\sigma(\sigma_8)$	$\sigma(w_0)$	$\sigma(w_a)$	$\sigma(f_{NL})$
Lensing	0.035 (0.014)	0.012 (0.0016)	0.011 (0.0081)	0.10 (0.011)	0.033 (0.0040)	0.017 (0.011)	0.12 (0.087)	0.55 (0.34)	37 (19)
2D clustering photometric	0.064 (0.037)	0.0085 (0.0043)	0.050 (0.026)	0.12 (0.033)	0.027 (0.0036)	0.042 (0.028)	0.51 (0.29)	1.6 (0.94)	3.2 (3.0)
Lensing + 2D photometric	0.014 (0.0069)	0.0035 (0.000 77)	0.0034 (0.0032)	0.033 (0.0050)	0.011 (0.0030)	0.0035 (0.0032)	0.037 (0.035)	0.19 (0.13)	2.8 (2.6)
Simple sum	0.024 (0.011)	0.0040 (0.0012)	0.0079 (0.0066)	0.038 (0.0089)	0.011 (0.0031)	0.0075 (0.0065)	0.089 (0.074)	0.39 (0.25)	2.8 (2.7)

Table 7. Forecasted results for non-Gaussianity for the different configurations. Constraints from galaxy clustering weaken in the orthogonal and equilateral cases, as the scale dependence of the bias is reduced. For the lensing (and thus also combined) cases, the constraints remain strong, as the non-Gaussian mass function is in all configurations expected to deviate at similar levels from the Gaussian model.

Euclid data (with *Planck* priors)	$\sigma(f_{NL})$ Local	Orthogonal	Equilateral
Lensing	37 (19)	4.9 (2.4)	17 (9.2)
2D clustering, photo-z	3.2 (3.0)	28 (8.1)	100 (30)
2D clustering, spectroscopic	6.7 (6.2)	50 (10)	150 (35)
3D clustering + shot	4.4 (4.2)	56 (12)	220 (37)
Lensing + 2D photometric	2.8 (2.6)	2.9 (1.6)	12 (6.5)
Lensing + 2D spectroscopic	4.9 (4.6)	3.4 (1.9)	11 (7.0)

simulations the predictions of the non-linear regime for the orthogonal and equilateral cases. This means that the constraints from lensing and its combinations in these two cases are necessarily less reliable than both the local case and the constraints from the scale-dependent bias, whose predictions are well tested and understood.

Finally, we can compare Tables 2, 3 and 4 to see how much the errors on the other parameters change. Here we can see that when the constraint on f_{NL} degrades greatly, such as in the galaxy clustering probes alone, the constraints on some other parameters also degrade significantly, especially σ_8. However the combined lensing+clustering result stays largely unchanged.

Our constraints on f_{NL} are weaker than what was found by Fedeli et al. (2011) for the 3D case only, as these authors calculated forecasts by varying only σ_8 and f_{NL} while fixing the rest of the parameters to ΛCDM and also used a more optimistic cut-off than ours for the non-linear regime.

6.4 Dependence on the fiducial

We then study the dependence of our forecasts on the choice of the fiducial values for the parameters, in particular extending the analysis to $f_{NL} \neq 0$ (in the case of local PNG only), or $w_0 \neq -1$. For this purpose, we run our full analysis for two non-standard cases, where the fiducials have the value of one parameter altered to $w_0 = -0.95$ and $f_{NL} = 30$, respectively.

6.4.1 Fiducial with dark energy

In the case of dark energy, we calculated the forecasted errors changing the value of the fiducial equation of state from $w_0 = -1$ to -0.95. We obtain that the constraints do not change significantly, as we can see in Table 8; we can conclude that the model with dark energy of $w_0 = -0.95$ would be effectively detected by *Euclid*, even without the need of *Planck* CMB priors.

6.4.2 Non-Gaussian fiducial

We chose to look at a local non-Gaussian model with $f_{NL} = 30$, as this is near the peak of the current posterior probability distribution given current data. We found that setting the fiducial to such a model degrades to some extent the accuracy of the errors. For the combined case (lensing + 2D photometric galaxy clustering), the uncertainty on f_{NL} degrades from $\sigma(f_{NL}) = 2.6$ to $\sigma(f_{NL}) = 4.1$ when *Planck* priors are included. We can see the results summarized in Table 9,

Table 8. Forecasted results for non-Gaussianity as a function of the fiducial value of w_0, without and with *Planck* priors. These results show that a *Euclid*-like survey should be able to distinguish between a cosmological constant and a different dark-energy model.

Euclid data (including *Planck* priors)	$\sigma(w_0)$ $w_0 = -1.0$	$w_0 = -0.95$
Lensing	0.12 (0.087)	0.11 (0.079)
2D clustering, photo-z	0.51 (0.29)	0.49 (0.29)
2D clustering, spectroscopic	0.12 (0.092)	0.091 (0.084)
3D clustering + shot noise	0.13 (0.031)	0.12 (0.031)
Lensing + 2D photometric	0.037 (0.035)	0.035 (0.033)
Lensing + 2D spectroscopic	0.052 (0.047)	0.040 (0.039)

Table 9. Forecasted results for non-Gaussianity as a function of the fiducial value of $f_{\rm NL}$ without and with *Planck* priors: detection versus upper limit scenarios. These results show that a *Euclid*-like survey will be able to either detect models with PNG down to the level of $f_{\rm NL} \sim 10$, or to find strict constraints around a Gaussian model.

Euclid data (including *Planck* priors)	$\sigma(f_{\rm NL})$ $f_{\rm NL} = 0$	$f_{\rm NL} = 30$
Lensing	37 (19)	38 (18)
2D clustering, photo-z	3.2 (3.0)	4.9 (4.3)
2D clustering, spectroscopic	6.7 (6.2)	8.1 (7.3)
3D clustering + shot noise	4.4 (4.2)	5.0 (4.7)
Lensing + 2D photometric	2.8 (2.6)	5.1 (4.1)
Lensing + 2D spectroscopic	4.9 (4.6)	5.4 (5.0)

from which we can conclude that this model would be detected with high significance by *Euclid*.

6.4.3 Galaxy biasing

The scale-dependent non-Gaussian correction to the galaxy linear bias is proportional to the Lagrangian bias $b_{\rm g} - 1$. Therefore, in the limit where the scale-independent part is $b_{\rm g}(z) \equiv 1$, the effect of PNG will vanish making it impossible to constrain $f_{\rm NL}$ from galaxy clustering on large scales. On the other hand, it is interesting to note that in the presence of highly biased tracers, such as luminous red galaxies or quasars, the effect – and the constraints – are maximized.

We look at how much the constraints are modified if we alter our fiducial bias by a fraction $f_{\rm b}$, i.e. we take

$$b_{\rm g}^{\rm fid}(z) = 1 + f_{\rm b}\,(\sqrt{1+z} - 1), \tag{47}$$

and we consider two cases with $f_{\rm b} = 1 \pm 0.25$. We can see in Table 10 that the constraints on $f_{\rm NL}$ are indeed strongly dependent on this choice, as expected, except obviously for the lensing case where there is no biasing.

6.5 Scale-dependent non-Gaussianity

The local shape of non-Gaussianity can be obtained e.g. when multiple scalar fields give a contribution to the curvature perturbations. However, in many models of inflation, non-Gaussianity is generated in a scale-dependent way. For this reason, a new parameter has been introduced by Chen (2005) in the equilateral case and by Byrnes et al. (2010) in the local case: the spectral index of non-Gaussianity,

Table 10. Marginalized forecasted results from *Euclid* for non-Gaussianity as a function of the fiducial value of the bias $b_{\rm g}^{\rm fid}(z)$, without and with *Planck* priors.

Euclid data (including *Planck* priors)	$\sigma(f_{\rm NL})$ $f_{\rm b} = 0.75$	$f_{\rm b} = 1$	$f_{\rm b} = 1.25$
Lensing	37 (19)	37 (19)	37 (19)
2D clustering, photo-z	3.8 (3.6)	3.2 (3.0)	2.8 (2.6)
2D clustering, spectroscopic	8.6 (7.9)	6.7 (6.2)	5.8 (5.6)
3D clustering + shot noise	5.7 (5.5)	4.4 (4.2)	3.6 (3.5)
Lensing + 2D photometric	4.2 (3.7)	2.8 (2.6)	2.4 (2.3)
Lensing + 2D spectroscopic	6.3 (5.7)	4.9 (4.6)	4.1 (3.9)

Table 11. Forecasted results from *Euclid*-like data for scale-dependent non-Gaussianity and degraded constraints on the scale-independent part $\bar{f}_{\rm NL}$. The fiducial model has here $\bar{f}_{\rm NL} = 30$.

Euclid data	$\sigma(\bar{f}_{\rm NL})$	$\sigma(n_{f_{\rm NL}})$
Lensing	68 (58)	0.66 (0.59)
2D clustering, photo-z	14 (9.6)	0.38 (0.26)
2D clustering, spectroscopic	23 (14)	0.64 (0.38)
3D clustering + shot noise	10 (7.6)	0.28 (0.21)
Lensing + 2D photometric	5.3 (4.3)	0.18 (0.14)
Lensing + 2D spectroscopic	6.6 (5.2)	0.17 (0.12)

$n_{f_{\rm NL}}$, defined so that the effective $f_{\rm NL}(k)$ is written as

$$f_{\rm NL}(k) = \bar{f}_{\rm NL} \left(\frac{k}{k_{\rm piv}}\right)^{n_{f_{\rm NL}}}, \tag{48}$$

where $\bar{f}_{\rm NL}$ is the value of $f_{\rm NL}$ at the pivot scale $k_{\rm piv}$. We choose the pivot scale to be near the logarithmic centre of the expected data (Cortes, Liddle & Mukherjee 2007) and fix it to $k_{\rm piv} = 0.02\,h\,{\rm Mpc}^{-1}$.

For the simplest case of local non-Gaussianity, in which a single field is responsible for generating the curvature perturbation such as the curvaton scenario (Huang 2010, 2011; Byrnes et al. 2011), it is possible to take the scale dependence of $f_{\rm NL}$ out of the integrals and simply apply it to the bias variation $\Delta b(k)$ (Desjacques et al. 2011; Shandera, Dalal & Huterer 2011). In this case, by adding $n_{f_{\rm NL}}$ to our parameter array, and fixing the fiducial $\bar{f}_{\rm NL} = 30$, we obtain the results of Table 11. The best constraint comes from combining lensing, clustering, and the *Planck* prior, giving $\sigma(n_{f_{\rm NL}}) \simeq 0.12$. The constraints on $f_{\rm NL}$ get in this case weakened. These constraints are similar to the forecasted limits from the *Planck* CMB bispectrum (Sefusatti et al. 2009), and fully independent.

6.6 Smooth bias parametrization

The constraints obtained so far are very conservative in the treatment of the bias. By assigning to every redshift bin an independent nuisance bias parameter, we are actually allowing for much more freedom than it is physically reasonable to expect. In particular, we have good reasons to expect the galaxy bias to be a smooth function of redshift. Therefore, we propose an alternative, less conservative, parametrization, in which the bias is assumed to be a polynomial function of redshift. In the following, we will use a third-order polynomial:

$$b(z) = b_0 + b_1(z-1) + b_2(z-1)^2 + b_3(z-1)^3\,, \tag{49}$$

with fiducial values $b_0 = \sqrt{2}$, $b_1 = (2\sqrt{2})^{-1}$, $b_2 = -(16\sqrt{2})^{-1}$, $b_3 = (64\sqrt{2})^{-1}$. The expansion in $(z-1)$ is required

if we want the bias to approximate $\sqrt{1+z}$ over the interval $0 < z < 2$. Note that this expansion converges only in this interval, and thus it would have to be modified for a broader redshift range. We have also checked that by expanding up to the third order we are able to approximate the $\sqrt{1+z}$ function to better than 1 per cent in this range.

When including the shot noise marginalization in this case, we also parametrize it as a third-order polynomial instead of leaving it as a completely free set of independent nuisance parameters for each bin, for the same reasons of requiring that physical quantities should have a smooth evolution in redshift. The forecasted errors shrink in this case however only marginally, as we can see in Table 12.

Finally, we evaluated how well could galaxy bias be measured with these observables. Referring to the smooth polynomial model of equation (49), we calculated the marginalized errors on its coefficients, which can be seen for the different *Euclid* probes in Table 13. Again, the combination of clustering+lensing gives the strictest results, since in this case all the other parameters are much better constrained, yielding an accuracy on the bias at the level of 1 per cent.

6.7 Dependence on the cut-off scales

6.7.1 Small scales

Here we study how the constraints change if we vary the maximum mode $k_{\rm max}$, $l_{\rm max}$ used in the forecasts. Since the constraints on each individual parameter will not in general vary monotonically, we will study the most interesting quantity, the figure of merit (FoM) **ssF** of the survey with respect to a set of N parameters $\boldsymbol{\vartheta}$, defined similarly to Albrecht & Bernstein (2007) as

$$\mathbf{F}(\boldsymbol{\vartheta}) = \frac{1}{\prod_{i=1}^{N} \tilde{\sigma}(\boldsymbol{\vartheta}_i)}\,, \tag{50}$$

where each $\tilde{\sigma}(\boldsymbol{\vartheta}_i)$ is the width of the error ellipsoid along the axis defined by the ith eigenvector of the Fisher matrix. In other words, it is easy to obtain their products by taking the determinant of the Fisher matrix as

$$\mathbf{F}(\boldsymbol{\vartheta}) = \sqrt{|\boldsymbol{F}(\boldsymbol{\vartheta})|}. \tag{51}$$

We then study the total FoM, the FoM obtained for the dark-energy parameters only, and for the dark-energy+$f_{\rm NL}$ parameters. We can see the result in Fig. 7, where we report the evolution of the FoM for lensing and the 3D power spectrum as a function of $l_{\rm max}$ and $k_{\rm max}$, respectively. In the 3D case, we evolve the scale cut-off with redshift as explained in the previous sections, by imposing that the variance of the density field stays constant in all redshift bins. In addition, to avoid considering exceedingly small scales, we impose a further condition that in every bin we will only consider scales $k \leq k_{\rm max}^{\rm abs} = 2k_{\rm max}(z=0)$.

As expected, the results are strongly dependent on the choice of the minimum scales used. For the case of galaxy clustering, we can see that we could still largely improve the constraints by extending the analysis deeper into the non-linear regime, with the use of more advanced biasing models reliable on smaller scales. On the other hand, we see that for weak lensing we have already included most of the signal with our choice of $l_{\rm max}$. It can be seen that more conservative choices would still yield a very similar result; for example by choosing more conservatively $l_{\rm max} = 5000$, we have that the marginalized error on $f_{\rm NL}$ with *Planck* priors degrades from 19 to 22 only, and the total FoM degrades by only a factor of 3.

Table 12. Expected marginalized errors for *Euclid*-like data for all parameters, using a parametric form of the galaxy bias. The numbers within parentheses include the forecasted priors from the *Planck* CMB temperature power spectrum. For 3D, includes marginalization over FoG.

Euclid data	$\sigma(\Omega_\Lambda)$	$\sigma(\Omega_{\rm b})$	$\sigma(\Omega_{\rm m})$	$\sigma(h)$	$\sigma(n)$	$\sigma(\sigma_8)$	$\sigma(w_0)$	$\sigma(w_{\rm a})$	$\sigma(f_{\rm NL})$
2D clustering photometric	0.054 (0.029)	0.0063 (0.0037)	0.032 (0.022)	0.099 (0.028)	0.027 (0.0036)	0.030 (0.024)	0.27 (0.22)	0.80 (0.67)	3.0 (3.0)
2D clustering spectroscopic	0.076 (0.013)	0.0062 (0.0018)	0.013 (0.0093)	0.089 (0.013)	0.047 (0.0039)	0.031 (0.014)	0.085 (0.052)	0.67 (0.28)	7.0 (6.6)
3D clustering + shot	0.078 (0.0080)	0.0043 (0.00091)	0.019 (0.0030)	0.026 (0.0056)	0.020 (0.0036)	0.0060 (0.0042)	0.13 (0.031)	0.67 (0.10)	4.4 (4.1)
Lensing + 2D photometric	0.014 (0.0068)	0.0035 (0.000 75)	0.0033 (0.0031)	0.033 (0.0049)	0.011 (0.0030)	0.0034 (0.0032)	0.036 (0.035)	0.18 (0.13)	2.8 (2.6)
Lensing + 2D spectroscopic	0.016 (0.0071)	0.0049 (0.000 81)	0.0038 (0.0034)	0.043 (0.0053)	0.014 (0.0035)	0.0041 (0.0038)	0.039 (0.035)	0.21 (0.13)	4.9 (4.6)

Table 13. Forecasted results from *Euclid* for the measurement of the galactic bias. A third-order polynomial expansion is assumed.

Euclid data	$\sigma(b_0)$	$\sigma(b_1)$	$\sigma(b_2)$	$\sigma(b_3)$
2D clustering, photo-z	0.038	0.031	0.016	0.081
2D clustering, spectroscopic	0.067	0.038	0.017	0.026
3D clusturing + shot noise	0.015	0.015	0.011	0.012
Lensing + 2D photometric	0.0048	0.0047	0.0046	0.0045
Lensing + 2D spectroscopic	0.0066	0.0076	0.012	0.024

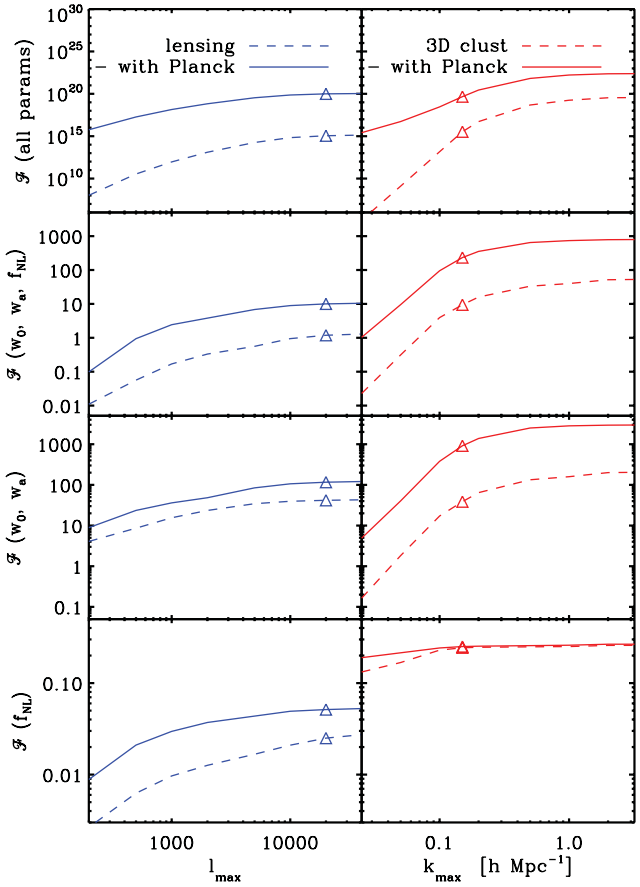

Figure 7. Dependence of the *Euclid*-like constraints on the small-scale cuts. In the top panels we show the evolution of the full FoM as a function of the minimum scale considered in the analysis: $l_{\rm max}$ for the 2D part (left) and $k_{\rm max}$ at $z = 0$ for the 3D part (right). In the following rows, we show the corresponding FoM relative to different combinations of few selected parameters: w_0, $w_{\rm a}$, $f_{\rm NL}$. The dashed lines denote *Euclid* only, while the solid lines include the *Planck* CMB priors (temperature power spectrum only). The triangles mark the scale which was chosen in the main analysis.

6.7.2 *Large scales*

We then also look at the effect of changing the large-scale cut-offs $k_{\rm min}$, $l_{\rm min}$. Due to the larger error bars, this is largely unimportant for most cosmological parameters, with the important exception of $f_{\rm NL}$: due to the form of the scale-dependent bias, the large scales are the region where this parameter is most constrained, and thus the forecasted FoM on $f_{\rm NL}$ is strongly dependent on this choice, as we can see in Fig. 8 for both the 2D and 3D cases. However, due to the limited size of the surveys, we decide that a conservative choice can hardly push to scales larger than $k_{\rm min} = 10^{-3}\,h\,{\rm Mpc}^{-1}$ and $l_{\rm min} = 5$, which we therefore adopt. A further reason to discard larger

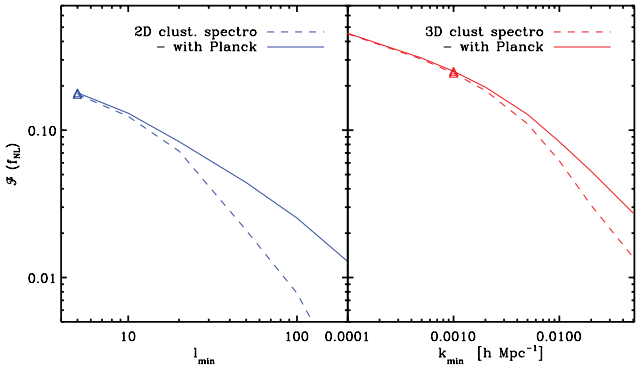

Figure 8. Dependence of the *Euclid* constraints on the large-scale cuts for the 2D and 3D cases. The FoM relative to the $f_{\rm NL}$ parameter only is shown. The value chosen in the main analysis, $k_{\rm min} = 10^{-3}\,h\,{\rm Mpc}^{-1}$, is marked with a triangle. The FoM for the remaining cosmological parameters is largely independent from this choice.

scales is to avoid the regime where general relativistic corrections become important, as described e.g. by Yoo (2010).

6.8 Effect of the halo model inaccuracies

As we have seen in Section 3, the accuracy of the halo model at small scales is only in the range of 10 per cent. For this reason we have tested how severely this impacts on the Fisher matrix forecasts derived using this model. For this purpose, we have calculated two sets of Fisher matrices for the *Euclid* spectroscopic survey, using the 3D power spectrum observable. For the first matrix of each set, we have derived the power spectra used in the halo model; for the second, we have used the approach by Smith et al. (2007). We can see in Fig. 9 the results using scales up to $k_{\rm max} = 0.15$ (as used in the main results) and 0.3 at $z = 0$, respectively, from which we can conclude that the our results are rather robust in the transition to the non-linear regime.

For weak lensing, where smaller scales are considered, the discrepancy between the two models for the non-linear matter power spectrum grows larger. Fig. 1 shows that the model by Smith et al. (2007) deviates from N-body simulations by more than 20 per cent for $k > 2\,h\,{\rm Mpc}^{-1}$, whereas the halo model is significantly more accurate on these scales. Consistently, previous studies have shown that Smith et al. (2007) underestimate the convergence power spectrum by more than 20 per cent at $l > 1000$, while the halo model is reliable up to $l \sim 50\,000$ (see fig. 9 in Hilbert et al. 2009). For this reason we consider only the halo model in our weak-lensing analysis.

6.9 Redshift binning

A further question which may be asked is how important is the choice of the redshift tomography used, i.e. how much would the results change if we used a different binning. We can see in Fig. 10 that this choice is not critical for the lensing and the 3D power spectrum, as in these cases the results have largely converged when we take more than a few bins. On the other hand, for the projected 2D spectra this choice is very important: due to the increasing number of cross-correlations between the bins, the signal-to-noise ratio increases significantly up to a rather high number of bins. For computational reasons, we have decided to carry our analysis using 12 redshift bins, where most cases are quite close to saturation.

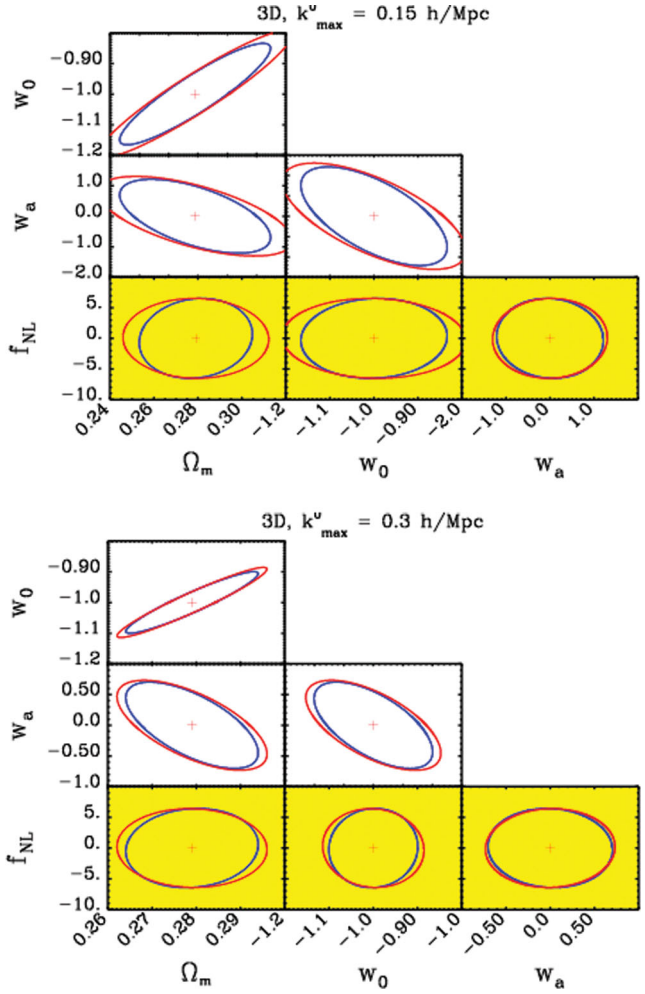

Figure 9. Effect of the uncertainties in the modelling of the non-linear regime on to the Fisher matrix forecasts for the 3D power spectrum in the local configuration. The ellipses in blue and red denote here the forecasted posteriors obtained using the halo model and the power spectra from the method by Smith et al. (2007), respectively. Only scales up to $k_{\rm max} = 0.15$ and $0.3\,h\,{\rm Mpc}^{-1}$ at $z = 0$ were used in the panels.

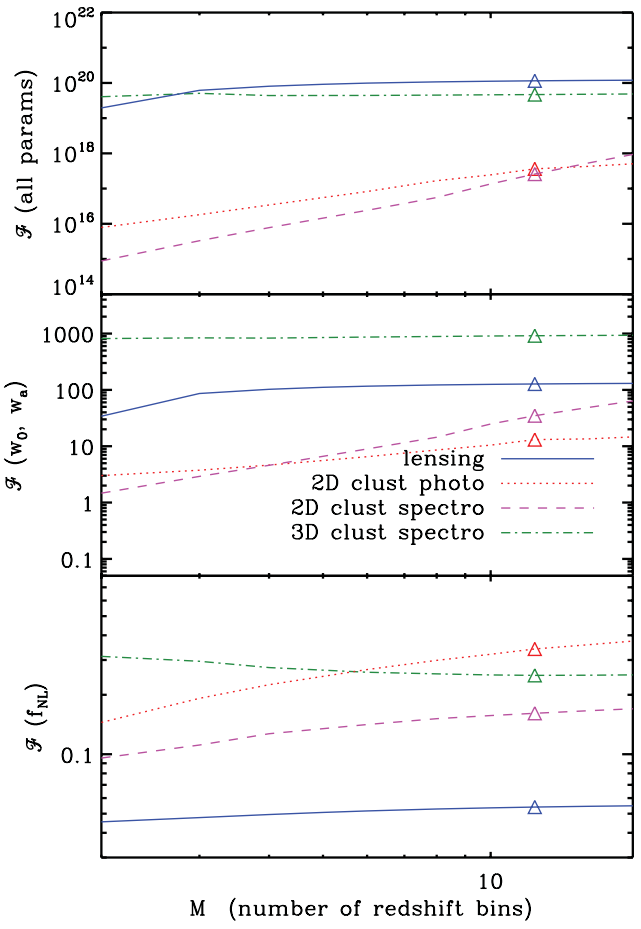

Figure 10. Dependence of the *Euclid* constraints on the redshift tomography. We show the evolution of the FoM as a function of the number of redshift bins considered in the analysis, for all the parameters (top), for the dark-energy parameters only (middle) and for $f_{\rm NL}$ (bottom panel). All these results include the *Planck* priors, and the choice used in the main analysis ($M = 12$) is marked by a triangle. We can see that the results for lensing and the 3D power spectrum have already converged once more than a few bins are used, while for the projected 2D cases the increase in the number of bins brings in much additional signal.

6.10 *Euclid* Red Book

The detailed specifications of the *Euclid* mission are still evolving. While the current study was performed based on the so-called Yellow Book (Laureijs 2009), the specifications have since evolved to the Red Book (Laureijs et al. 2011). In this section we describe how our results change as a consequence of this update.

We show in Table 14 the parameters which have changed from the Yellow to the Red Book; for the latter, in many cases two values are given: minimum requirement and goal. As can be seen, the 'goal' specification is generally the same as the Yellow Book value, while the 'requirement' is less ambitious. The main exception is the galaxy density in the spectroscopic case, which has increased due to a deeper flux limit. Table 15 reports the changes in the forecasted constraints for the Red Book 'requirement' specifications for the local PNG case. The Red Book 'goal' specifications are very similar to the Yellow Book results presented above. We can see that in general the constraints degrade for the photometric survey

Table 14. Changes in the specifications for the *Euclid* Red Book (Laureijs et al. 2011). The parameters which are not shown are unchanged from Table 1.

Red Book update		*Euclid* photometric		*Euclid* spectroscopic	
Parameter	Description	Required	Goal	Required	Goal
$\sigma_z(z)/(1+z)$	Redshift uncertainty	0.05	0.03	0.001	0.001
$\bar{z}$	Median redshift	0.8	1.0	1.0	1.1
n	Galaxy density	30 arcmin^{-2}	40 arcmin^{-2}	1.20 arcmin^{-2}	1.20 arcmin^{-2}
A	Surveyed area	15 000 square degrees	20 000 square degrees	15 000 square degrees	20 000 square degrees
$dN/dz(z)$	Galaxy distribution	Smail et al.		Geach et al.	

Table 15. Expected marginalized errors from *Euclid* for the Red Book specifications (minimum requirements) for the local non-Gaussianity case. The numbers within parentheses include the forecasted priors from the *Planck* CMB temperature power spectrum. The 3D clustering includes marginalization over the shot noise parameters in each redshift bin. For the spectroscopic part of the survey, we show both the results obtained using our standard form for fiducial bias $b(z) \propto \sqrt{1+z}$ and the results from the tabulated bias by Orsi et al. (2010) (in italic).

Red Book	$\sigma(\Omega_\Lambda)$	$\sigma(\Omega_b)$	$\sigma(\Omega_m)$	$\sigma(h)$	$\sigma(n)$	$\sigma(\sigma_8)$	$\sigma(w_0)$	$\sigma(w_a)$	$\sigma(f_{NL})$
Lensing	0.078 (0.017)	0.019 (0.0019)	0.016 (0.011)	0.16 (0.014)	0.063 (0.0041)	0.030 (0.015)	0.17 (0.12)	1.0 (0.47)	73 (27)
2D clustering photometric	0.095 (0.036)	0.0087 (0.0040)	0.050 (0.024)	0.15 (0.030)	0.043 (0.0039)	0.034 (0.026)	0.50 (0.31)	1.9 (1.2)	5.8 (5.5)
2D clustering spectroscopic	0.076 (0.015)	0.0065 (0.0021)	0.015 (0.011)	0.099 (0.015)	0.050 (0.0038)	0.031 (0.015)	0.090 (0.058)	0.54 (0.28)	6.4 (5.9)
–, *tab bias*	*0.082 (0.016)*	*0.0069 (0.0023)*	*0.017 (0.012)*	*0.11 (0.016)*	*0.056 (0.0039)*	*0.036 (0.017)*	*0.10 (0.063)*	*0.59 (0.30)*	*8.8 (8.0)*
3D + shot noise	0.067 (0.0083)	0.0048 (0.0010)	0.023 (0.0037)	0.027 (0.0064)	0.033 (0.0035)	0.042 (0.017)	0.13 (0.029)	0.59 (0.092)	4.1 (4.0)
–, *tab bias*	*0.071 (0.0084)*	*0.0051 (0.0010)*	*0.025 (0.0038)*	*0.029 (0.0065)*	*0.036 (0.0035)*	*0.048 (0.020)*	*0.14 (0.031)*	*0.63 (0.098)*	*6.1 (5.8)*
Lensing + 2D photometric	0.030 (0.0087)	0.0056 (0.000 97)	0.0048 (0.0044)	0.050 (0.0065)	0.016 (0.0036)	0.0052 (0.0048)	0.054 (0.052)	0.32 (0.20)	4.7 (4.5)
Lensing + 2D spectroscopic	0.030 (0.0077)	0.0052 (0.000 89)	0.0051 (0.0037)	0.048 (0.0058)	0.016 (0.0036)	0.0054 (0.0045)	0.051 (0.035)	0.35 (0.15)	5.7 (5.3)
–, *tab bias*	*0.031 (0.0080)*	*0.0054 (0.000 92)*	*0.0054 (0.0040)*	*0.051 (0.0061)*	*0.018 (0.0037)*	*0.0060 (0.0051)*	*0.055 (0.039)*	*0.38 (0.16)*	*7.9 (7.3)*

and slightly improve for the spectroscopic part, mainly due to the increased number density in this latter case.

In addition, we also show the results obtained using a fiducial bias derived from semianalytic models of galaxy formation (Orsi et al. 2010), which are available for the spectroscopic case only. In this case the constraints on f_{NL} degrade further, as this tabulated bias is approximately 15 per cent lower than the model $b(z) = \sqrt{1+z}$ which we use in our main results.

7 CONCLUSIONS

In this work we have studied to what accuracy will the two-point statistics of future surveys of the LSS determine the cosmological parameters, with a particular focus on the non-Gaussianity parameter f_{NL}, considering the most relevant local, equilateral and orthogonal bispectrum configurations. We have performed a Fisher matrix analysis using the specifications of the upcoming DES and a *Euclid*-like survey based on the *Euclid* Assessment Study Report, using both its spectroscopic and photometric parts. We have chosen these surveys as examples of the future DETF stages III and IV, and the results are likely comparable with other surveys in the same class, such as in particular the planned American mission *WFIRST*.

We have combined all the relevant data sets, including their covariances. In particular, we have considered the projected 2D galaxy power spectrum, always including curvature and using the exact calculation, discarding the Limber approximation on large scales, as this introduces inaccuracies which are important in the presence of a scale-dependent galaxy bias from PNG.

We obtained that the strictest constraints on f_{NL} are expected from the combination of weak lensing and photometric galaxy clustering; in this case we find for the local case $\sigma(f_{NL}) \simeq 3$ for *Euclid* and $\simeq 8$ for the DES, when also including priors from the temperature power spectra of the *Planck* CMB mission. In the cases of orthogonal and equilateral configurations, the constraints from galaxy clustering degrade greatly, due to the reduced scale dependence of the galaxy bias, while the constraints from weak lensing remain at a similar level. Finally, the constraint on scale-dependent non-Gaussianity is for the local case $\sigma(n_{f_{NL}}) = 0.12$ when the fiducial scale-independent part is $\bar{f}_{NL} = 30$. The level of these constraints is comparable to the expectations from the *Planck* CMB bispectrum, and fully independent (Komatsu & Spergel 2001; Sefusatti et al. 2009).

We have also studied the effect of updating the description of *Euclid* to the latest Red Book specifications; in this case using the Red Book 'goal' parameters leaves the forecasts largely unchanged, while we found that using the 'requirement' parameters degrades the constraints on PNG to $\sigma(f_{NL}) \simeq 5$.

On one hand, further extensions of this approach, which will be useful not only for forecasting but also for the likelihood analysis of the upcoming real data, will include expanding the total covariance of the two-point statistics to include observations from clusters of galaxies and the correlations with the CMB; on the other hand, this approach will be extended to the inclusion of the three- and four-point statistics, which will be instrumental in the search for the higher order g_{NL} and τ_{NL} inflationary parameters.

ACKNOWLEDGMENTS

We acknowledge Jochen Weller for supplying us with the CMB Fisher matrix code and Christian Byrnes, Martin Kilbinger, Fabian Schmidt and Masahiro Takada for useful discussions. We thank the Euclid collaboration, Yannick Mellier and in particular the theory

working group for stimulating discussions. TG acknowledges support from the Alexander von Humboldt Foundation. JC and AP acknowledge support from the Swiss Science National Foundation. CP acknowledges support from the German Research Association (DFG) through the Transregional Collaborative Research Center 'The Dark Universe' (TRR33).

REFERENCES

Abbott L. F., Schaefer R. K., 1986, ApJ, 308, 546
Abbott T. et al., 2005, preprint (astro-ph/0510346)
Afshordi N., Tolley A. J., 2008, Phys. Rev. D, 78, 123507
Aihara H. et al., 2011, ApJS, 193, 29
Albrecht A., Bernstein G., 2007, Phys. Rev. D, 75, 103003
Albrecht A. et al., 2006, preprint (astro-ph/0609591)
Albrecht A. et al., 2009, preprint (arXiv:0901.0721)
Amara A., Refregier A., 2007, MNRAS, 381, 1018
Babich D., Creminelli P., Zaldarriaga M., 2004, J. Cosmol. Astropart. Phys., 0408, 009
Bardeen J. M., Bond J. R., Kaiser N., Szalay A. S., 1986, ApJ, 304, 15
Bartelmann M., Schneider P., 2001, Phys. Rep., 340, 291
Bernstein G., Huterer D., 2010, MNRAS, 401, 1399
Bryan G., Norman M., 1998, ApJ, 495, 80
Bullock J. S., Kolatt T. S., Sigad Y., Somerville R. S., Kravtsov A. V., Klypin A. A., Primack J. R., Dekel A., 2001, MNRAS, 321, 559
Byrnes C. T., Choi K.-Y., 2010, Adv. Astron., 2010, 724525
Byrnes C. T., Nurmi S., Tasinato G., Wands D., 2010, J. Cosmol. Astropart. Phys., 1002, 034
Byrnes C. T., Enqvist K., Nurmi S., Takahashi T., 2011, J. Cosmol. Astropart. Phys., 1111, 011
Carbone C., Verde L., Matarrese S., 2008, ApJ, 684, L1
Carbone C., Mena O., Verde L., 2010, J. Cosmol. Astropart. Phys., 1007, 020
Catelan P., Lucchin F., Matarrese S., Porciani C., 1998, MNRAS, 297, 692
Chen X., 2005, Phys. Rev. D, 72, 123518
Chen X., 2010, Adv. Astron., 2010, 638979
Chevallier M., Polarski D., 2001, Int. J. Modern Phys. D, 10, 213
Cole S., Kaiser N., 1989, MNRAS, 237, 1127
Cooray A., Sheth R. K., 2002, Phys. Rep., 372, 1
Cortes M., Liddle A. R., Mukherjee P., 2007, Phys. Rev. D, 75, 083520
Dalal N., Dore O., Huterer D., Shirokov A., 2008, Phys. Rev. D, 77, 123514
D'Aloisio A., Natarajan P., 2011, MNRAS, 415, 1913
Desjacques V., Seljak U., 2010a, Classical Quantum Gravity, 27, 124011
Desjacques V., Seljak U., 2010b, Adv. Astron., 2010, 908640
Desjacques V., Seljak U., Iliev I. T., 2009, MNRAS, 396, 85
Desjacques V., Jeong D., Schmidt F., 2011, Phys. Rev. D, 84, 061301
Fedeli C., Moscardini L., 2010, MNRAS, 405, 681
Fedeli C., Carbone C., Moscardini L., Cimatti A., 2011, MNRAS, 414, 1545
Feldman H. A., Kaiser N., Peacock J. A., 1994, ApJ, 426, 23
Fergusson J. R., Shellard E. P. S., 2009, Phys. Rev. D, 80, 043510
Frieman J., Turner M., Huterer D., 2008, ARA&A, 46, 385
Fu L. et al., 2008, A&A, 479, 9
Geach J. E. et al., 2010, MNRAS, 402, 1330
Giannantonio T., Porciani C., 2010, Phys. Rev. D, 81, 063530
Giannantonio T., Scranton R., Crittenden R. G., Nichol R. C., Boughn S. P., Myers A. D., Richards G. T., 2008, Phys. Rev. D, 77, 123520
Grossi M., Verde L., Carbone C., Dolag K., Branchini E., Iannuzzi F., Matarrese S., Moscardini L., 2009, MNRAS, 398, 321
Hilbert S., Hartlap J., White S. D. M., Schneider P., 2009, A&A, 499, 31
Hoyle B., Jimenez R., Verde L., 2011, Phys. Rev. D, 83, 103502
Hoyle B., Jimenez R., Verde L., Hotchkiss S., 2012, J. Cosmol. Astropart. Phys., 2, 9
Hu W., Jain B., 2004, Phys. Rev. D, 70, 043009
Huang Q.-G., 2010, J. Cosmol. Astropart. Phys., 1011, 026
Huang Q.-G., 2011, J. Cosmol. Astropart. Phys., 1102, E01
Jenkins A., Frenk C. S., White S. D. M., Colberg J. M., Cole S., Evrard A. E., Couchman H. M. P., Yoshida N., 2001, MNRAS, 321, 372
Jeong D., Komatsu E., 2009, ApJ, 703, 1230
Joachimi B., Bridle S. L., 2010, A&A, 523, A1
Kaiser N., 1987, MNRAS, 227, 1
Kessler R. et al., 2009, ApJS, 185, 32
Kirk D., Bridle S., Schneider M., 2010, MNRAS, 408, 1502
Kitching T. D., Taylor A. N., Heavens A. F., 2008, MNRAS, 389, 173
Klypin A., Trujillo-Gomez S., Primack J., 2011, ApJ, 740, 102
Koester B. et al., 2007, ApJ, 660, 239
Komatsu E., 2010, Classical Quantum Gravity, 27, 124010
Komatsu E., Spergel D. N., 2001, Phys. Rev. D, 63, 063002
Komatsu E. et al., 2009, in Astro2010: The Astronomy and Astrophysics Decadal Survey, Vol. 2010, p. 158
Komatsu E. et al., 2011, ApJS, 192, 18
Lam T. Y., Sheth R. K., 2009, MNRAS, 398, 2143
Lampeitl H. et al., 2009, MNRAS, 401, 2331
Larson D. et al., 2011, ApJS, 192, 16
Laureijs R., 2009, preprint (arXiv:0912.0914)
Laureijs R. et al., 2011, preprint (arXiv:1110.3193)
Lewis A., Challinor A., Lasenby A., 2000, ApJ, 538, 473
Liguori M., Sefusatti E., Fergusson J., Shellard E., 2010, Adv. Astron., 2010, 980523
Linde A. D., 2008, in Lemoine M., Martin J., Peter P., eds, Lect. Notes Phys. Vol. 738, Inflationary Cosmology. Springer, Berlin, p. 1
LoVerde M., Miller A., Shandera S., Verde L., 2008, J. Cosmol. Astropart. Phys., 0804, 014
Lucchin F., Matarrese S., 1988, ApJ, 330, 535
Lyth D. H., Liddle A. R., 2009, The Primordial Density Perturbation: Cosmology, Inflation and the Origin of Structure. Cambridge Univ. Press, Cambridge, p. 497
Ma C.-P., Fry J. N., 2000, ApJ, 543, 503
McClelland J., Silk J., 1977, ApJ, 217, 331
Maggiore M., Riotto A., 2010a, ApJ, 717, 526
Maggiore M., Riotto A., 2010b, MNRAS, 405, L1244
Magliocchetti M., Porciani C., 2003, MNRAS, 346, 186
Marian L., Hilbert S., Smith R. E., Schneider P., Desjacques V., 2011, ApJ, 728, L13
Matarrese S., Verde L., 2008, ApJ, 677, L77
Matarrese S., Verde L., Jimenez R., 2000, ApJ, 541, 10
Mo H. J., White S. D. M., 1996, MNRAS, 282, 347
Mortonson M. J., Hu W., Huterer D., 2011, Phys. Rev. D, 83, 023015
Namikawa T., Okamura T., Taruya A., 2011, Phys. Rev. D, 83, 123514
Navarro J. F., Frenk C. S., White S. D. M., 1996, ApJ, 462, 563
Navarro J. F., Frenk C. S., White S. D. M., 1997, ApJ, 490, 493
Orsi A., Baugh C. M., Lacey C. G., Cimatti A., Wang Y., Zamorani G., 2010, MNRAS, 405, 1006
Peacock J. A., Dodds S. J., 1994, MNRAS, 267, 1020
Percival W. J., Cole S., Eisenstein D. J., Nichol R. C., Peacock J. A., Pope A. C., Szalay A. S., 2007, MNRAS, 381, 1053
Pillepich A., Porciani C., Hahn O., 2010, MNRAS, 402, 191 (PPH08)
Pillepich A., Porciani C., Reiprich T. H., 2011, preprint (arXiv:1111.6587)
Press W. H., Schechter P., 1974, ApJ, 187, 425
Rassat A. et al., 2008, preprint (arXiv:0810.0003)
Refregier A., Amara A., Kitching T., Rassat A., 2011, A&A, 528, A33
Roth N., Porciani C., 2011, MNRAS, 415, 829
Robertson B. E., Kravtsov A. V., Tinker J., Zentner A. R., 2009, ApJ, 696, 636
Rudd D. H., Zentner A. R., Kravtsov A. V., 2008, ApJ, 672, 19
Sartoris B., Borgani S., Fedeli C., Matarrese S., Moscardini L., Rosati P., Weller J., 2010, MNRAS, 407, 2339
Sato M., Hamana T., Takahashi R., Takada M., Yoshida N., Matsubara T., Sugiyama N., 2009, ApJ, 701, 945
Schmidt F., Kamionkowski M., 2010, Phys. Rev. D, 82, 103002
Scoccimarro R., Sheth R. K., Hui L., Jain B., 2001, ApJ, 546, 20
Sefusatti E., Komatsu E., 2007, Phys. Rev. D, 76, 083004
Sefusatti E., Liguori M., Yadav A. P., Jackson M. G., Pajer E., 2009, J. Cosmol. Astropart. Phys., 0912, 022

Seljak U., 2000, MNRAS, 318, 203
Senatore L., Smith K. M., Zaldarriaga M., 2010, J. Cosmol. Astropart. Phys., 1001, 028
Seo H.-J., Eisenstein D. J., 2003, ApJ, 598, 720
Shandera S., Dalal N., Huterer D., 2011, J. Cosmol. Astropart. Phys., 1103, 017
Sheth R. K., Tormen G., 1999, MNRAS, 308, 119
Sheth R. K., Tormen G., 2002, MNRAS, 329, 61
Simpson F., James J., Heavens A. F., Heymans C., 2011, Phys. Rev. Lett., 107, 271301
Slosar A., Hirata C., Seljak U., Ho S., Padmanabhan N., 2008, J. Cosmol. Astropart. Phys., 0808, 031
Smail I., Ellis R. S., Fitchett M. J., 1994, MNRAS, 270, 245
Smith R. E., Scoccimarro R., Sheth R. K., 2007, Phys. Rev. D, 75, 063512
Smith R. E., Desjacques V., Marian L., 2011, Phys. Rev. D, 83, 043526
Song Y.-S., Percival W. J., 2009, J. Cosmol. Astropart. Phys., 0910, 004
Takada M., Bridle S., 2007, New J. Phys., 9, 446
Takada M., Jain B., 2009, MNRAS, 395, 2065
Taruya A., Koyama K., Matsubara T., 2008, Phys. Rev. D, 78, 123534
Tegmark M., 1997, Phys. Rev. Lett., 79, 3806
Tegmark M., Taylor A., Heavens A., 1997, ApJ, 480, 22
Tinker J. L., Weinberg D. H., Zheng Z., Zehavi I., 2005, ApJ, 631, 41
Tinker J. L., Kravtsov A. V., Klypin A., Abazajian K., Warren M. S., Yepes G., Gottlober S., Holz D. E., 2008, ApJ, 688, 709
Valageas P., 2010, A&A, 514, A46
van Daalen M. P., Schaye J., Booth C., Vecchia C. D., 2011, MNRAS, 415, 3649
Wagner C., Verde L., 2011, preprint (arXiv:1102.3229)
Wang Y. et al., 2010, MNRAS, 409, 737
Warren M. S., Abazajian K., Holz D. E., Teodoro L., 2006, ApJ, 646, 881
Xia J.-Q., Baccigalupi C., Matarrese S., Verde L., Viel M., 2011, J. Cosmol. Astropart. Phys., 8, 33
Yoo J., 2010, Phys. Rev. D, 82, 083508
Zheng Z., Tinker J. L., Weinberg D. H., Berlind A. A., 2002, ApJ, 575, 617

APPENDIX A: PRIMORDIAL NON-GAUSSIANITY AND THE LARGE-SCALE STRUCTURE

We review here the effects of PNG on the LSS, focusing on the results which are used in this analysis.

A1 Bispectra and skewness

The bispectrum of the Bardeen potential Φ in the local, equilateral and orthogonal cases is given by (Taruya et al. 2008; Schmidt & Kamionkowski 2010)

$$B_\Phi^{\rm loc}(k_1,k_2,k_3) \simeq 2 f_{\rm NL}\left[P_\Phi(k_1)P_\Phi(k_2) + 2\,\text{perms.}\right], \tag{A1}$$

$$\begin{aligned} B_\Phi^{\rm equ}(k_1,k_2,k_3) \simeq 6 f_{\rm NL}^{\rm eq}\Big\{ &-\left[P_\Phi(k_1)P_\Phi(k_2) + 2\,\text{perms.}\right] \\ &-2\left[P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3)\right]^{2/3} \\ &+\left[P_\Phi^{1/3}(k_1)P_\Phi^{2/3}(k_2)P_\Phi(k_3) + 5\,\text{perms.}\right]\Big\}, \end{aligned} \tag{A2}$$

$$\begin{aligned} B_\Phi^{\rm ort}(k_1,k_2,k_3) \simeq 6 f_{\rm NL}^{\rm ort}\Big\{ &-3\left[P_\Phi(k_1)P_\Phi(k_2) + 2\,\text{perms.}\right] \\ &-8\left[P_\Phi(k_1)P_\Phi(k_2)P_\Phi(k_3)\right]^{2/3} \\ &+3\left[P_\Phi^{1/3}(k_1)P_\Phi^{2/3}(k_2)P_\Phi(k_3) + 5\,\text{perms.}\right]\Big\}. \end{aligned} \tag{A3}$$

It is interesting to compute the lowest moments of the corresponding linear density field δ: the variance and the skewness. The variance is defined as

$$\sigma^2(M) \equiv \left\langle \delta_M^2 \right\rangle = \frac{1}{2\pi^2}\int {\rm d}k\, k^2 P(k) F^2(k,M), \tag{A4}$$

where we have introduced the smoothed field $\delta_M(k) = \delta(k)F(k,M)$, and $F(k, M)$ is a filter function of mass resolution M. We use a top-hat function in real space.

The third momentum of the smoothed density field can be written in terms of the three-point function

$$\left\langle \delta_M^3 \right\rangle = \int \frac{{\rm d}^3k_1}{(2\pi)^3}\int \frac{{\rm d}^3k_2}{(2\pi)^3}\int \frac{{\rm d}^3k_3}{(2\pi)^3}\langle \delta_M(k_1)\delta_M(k_2)\delta_M(k_3)\rangle. \tag{A5}$$

By substituting the bispectra from equations (A1), (A2), (A3), and integrating the Dirac delta to give $k_3 = -k_1 - k_2$ so that we can define $k \equiv |k_1 + k_2|$, this expression can be simplified for the three cases extending the calculation by Desjacques, Seljak & Iliev (2009),

$$\begin{aligned} \left\langle \delta_M^3 \right\rangle^{\rm loc} = &\frac{f_{\rm NL}}{(2\pi^2)^2}\int_0^\infty {\rm d}k_1 k_1^2 \frac{P(k_1)}{\alpha(k_1)}\int_0^\infty {\rm d}k_2 k_2^2 \frac{P(k_2)}{\alpha(k_2)} \\ &\times \int_{-1}^1 {\rm d}\mu\, \alpha(k)\left[1 + 2\frac{P(k)\alpha^2(k_2)}{P(k_2)\alpha^2(k)}\right] \\ &\times F(k_1,M)\,F(k_2,M)\,F(k,M), \end{aligned} \tag{A6}$$

$$\begin{aligned} \left\langle \delta_M^3 \right\rangle^{\rm equ} = &\frac{3 f_{\rm NL}}{(2\pi^2)^2}\int_0^\infty {\rm d}k_1 k_1^2 \alpha(k_1)\int_0^\infty {\rm d}k_2 k_2^2 \alpha(k_2) \\ &\times \int_{-1}^1 {\rm d}\mu\, \alpha(k)\Bigg\{-\left[\frac{P(k_1)P(k_2)}{\alpha^2(k_1)\alpha^2(k_2)} + 2\,\text{perms.}\right] \\ &-2\left[\frac{P(k_1)P(k_2)P(k)}{\alpha^2(k_1)\alpha^2(k_2)\alpha^2(k)}\right]^{2/3} \\ &+\left[\frac{P^{1/3}(k_1)P^{2/3}(k_2)P(k)}{\alpha^{2/3}(k_1)\alpha^{4/3}(k_2)\alpha^2(k)} + 5\,\text{perms.}\right]\Bigg\} \\ &\times F(k_1,M)\,F(k_2,M)\,F(k,M), \end{aligned} \tag{A7}$$

$$\begin{aligned} \left\langle \delta_M^3 \right\rangle^{\rm ort} = &\frac{3 f_{\rm NL}}{(2\pi^2)^2}\int_0^\infty {\rm d}k_1 k_1^2 \alpha(k_1)\int_0^\infty {\rm d}k_2 k_2^2 \alpha(k_2) \\ &\times \int_{-1}^1 {\rm d}\mu\, \alpha(k)\Bigg\{-3\left[\frac{P(k_1)P(k_2)}{\alpha^2(k_1)\alpha^2(k_2)} + 2\,\text{perms.}\right] \\ &-8\left[\frac{P(k_1)P(k_2)P(k)}{\alpha^2(k_1)\alpha^2(k_2)\alpha^2(k)}\right]^{2/3} \\ &+3\left[\frac{P^{1/3}(k_1)P^{2/3}(k_2)P(k)}{\alpha^{2/3}(k_1)\alpha^{4/3}(k_2)\alpha^2(k)} + 5\,\text{perms.}\right]\Bigg\} \\ &\times F(k_1,M)\,F(k_2,M)\,F(k,M). \end{aligned} \tag{A8}$$

The skewness is finally defined for all cases as

$$S_3(M) \equiv \left\langle \delta_M^3 \right\rangle / \left\langle \delta_M^2 \right\rangle^2 = \left\langle \delta_M^3 \right\rangle / \sigma^4. \tag{A9}$$

A2 Mass function

The halo mass function $n(M, z)$ gives the number density of haloes of mass M at redshift z. Various analytical methods have been developed to compute this quantity starting from the statistical properties of Gaussian primordial perturbations, starting with Press & Schechter (1974). Fitting formulae that improve agreement with N-body simulations have also been presented (Sheth & Tormen 1999, 2002; Jenkins et al. 2001; Warren et al. 2006; Tinker et al. 2008; Pillepich et al. 2010).

The halo mass function is modified by the presence of PNG (Matarrese, Verde & Jimenez 2000; LoVerde et al. 2008; Maggiore & Riotto 2010a; Lam & Sheth 2009). Since analytical models are based on a set of simplistic assumptions, they most robustly predict the ratio between the mass function generated by non-Gaussian and Gaussian initial conditions with the same power spectrum [see Giannantonio & Porciani (2010) for a comparison of the different methods against N-body simulations].

For this reason, in our forecasts we use the fitting formula given by PPH08 for the Gaussian mass function and multiply it by the factor (LoVerde et al. 2008)

$$R_{\rm LV}\left(\frac{\delta_{\rm c}}{\sigma}, f_{\rm NL}\right) \equiv \left[1 + \frac{S_3\sigma^2}{6\sqrt{q}\,\delta_{\rm c}}\left(\frac{q^2\,\delta_{\rm c}^4}{\sigma^4} - \frac{2q\,\delta_{\rm c}^2}{\sigma^2} - 1\right) + \frac{{\rm d}S_3(\sigma)}{{\rm dln}\sigma^2}\frac{\sigma}{6\sqrt{q}\,\delta_{\rm c}}\left(\frac{q\,\delta_{\rm c}^2}{\sigma^2} - 1\right)\right], \tag{A10}$$

in the non-Gaussian case. Here $\delta_{\rm c} \simeq 1.686$ is the threshold for the collapse of a linear density perturbation and the corrective factor q, of order unity, is a heuristical correction which may be applied to the collapse threshold to improve the agreement with N-body simulations, and which we set to unity for the reasons given in Section 2. Note that PNG enters the expression for the halo mass function through the linear density skewness $S_3(M)$ which is linearly proportional to $f_{\rm NL}$.

A3 Halo bias

The halo linear bias factors may be obtained from the expression for the halo mass function using the peak-background split formalism (Bardeen et al. 1986; Cole & Kaiser 1989; Mo & White 1996; Catelan et al. 1998). In the Gaussian case, the bias is obtained by taking the logarithmic derivative of the mass function with respect to the collapse threshold. This gives a mass and redshift dependent bias coefficient.

Recent generalizations of the peak-background split to cases with PNG have shown that the corrections to the mass function generate two extra terms in the expression for the halo bias: one of them is a small addition to the Gaussian bias while the second one introduces a scale-dependent term (see equations (7) and (9) in the main text) (Dalal et al. 2008; Slosar et al. 2008; Matarrese & Verde 2008; Afshordi & Tolley 2008; Giannantonio & Porciani 2010; Schmidt & Kamionkowski 2010; Desjacques et al. 2011).

APPENDIX B: PLANCK–FISHER MATRIX

The Planck–Fisher matrix was kindly made available to us by Jochen Weller. It was calculated according to Albrecht et al. (2009) and Rassat et al. (2008) and used in Sartoris et al. (2010). As described in section 3.8 of Sartoris et al. (2010), it is computed adopting the parameterization $\boldsymbol{\vartheta} = (\omega_{\rm m}, \boldsymbol{\vartheta}_{\rm s}, \ln A_{\rm s}, \omega_{\rm b}, n_{\rm s}, \tau)$, with $A_{\rm s}$ the amplitude of primordial perturbations, $\boldsymbol{\vartheta}_{\rm s}$ the size of the sound horizon at the last-scattering surface, and τ the optical depth to Compton scattering which has been marginalized over. The calculation of the Fisher matrix is based on the conservative assumption that only the 143 Ghz channel is used for science analysis. In this case, the beam size is $\boldsymbol{\vartheta}_{\rm FWHM} = 7.1$ arcmin, and the temperature and polarization sensitivities are $\sigma_{\rm T} = 2.2\,\mu{\rm K\,K}^{-1}$ and $\sigma_{\rm P} = 4.2\,\mu{\rm K\,K}^{-1}$, respectively. In order to avoid polarization foregrounds, only multipoles with $l_{\rm min} = 30$ have been considered together with a sky fraction $f_{\rm sky} = 0.8$ to minimize the effects of galactic foregrounds. The transformation of the Fisher matrix for the parameter set ($\Omega_{\rm m}$, Ω_Λ, h, σ_8, $\Omega_{\rm b}$, w_0, $w_{\rm a}$, $n_{\rm s}$) is done as in Rassat et al. (2008). It is important to highlight that the Planck–Fisher matrix does not include any information on primordial non-Gaussianity, as CMB constraints on PNG can be calculated from higher-order statistics only (especially the bispectrum). So adding these priors will not improve the constraints on $f_{\rm NL}$.

This paper has been typeset from a TEX/LATEX file prepared by the author.